UNIVERSIDADE DE SÃO PAULO

Instituto de Matemática e Estatística

A precificação de opções para processos de mistura de brownianos

Herbert Kimura

Departamento de Estatística

Dissertação de Mestrado

Orientador: Prof. Dr. Vladimir Belitsky

Sumário

Chapter 1

Introdução

O raciocínio de Black e Scholes para precificação de opções é bem conhecido. No Capítulo 2, apresentamos os principais conceitos, definições e resultados da teoria de finanças que subsidiarão nosso estudo. Uma das hipóteses fundamentais para o raciocínio de Black e Scholes é a suposição de que o preço de ação segue um movimento browniano geométrico dado por

$$S_t = S_0 \exp\{\mu t + \sigma W_t\}, t \geq 0 \tag{1.1}$$

onde S_t designa o preço de ação no tempo t, μ e σ são duas constantes com a restrição $\sigma > 0$, e, finalmente, $W_t, t \geq 0$, é o movimento browniano padrão. A hipótese (1.1) implica, em particular, que os retornos do ativo-objeto têm distribuições normais:

$$\log\left\{\frac{S_{t+1}}{S_t}\right\} \sim \mathcal{N}(\mu, \sigma^2) \tag{1.2}$$

Porém, existem evidências empíricas [1] de que os dados amostrais não aderem bem à distribuição normal. Para grande parte dos ativos, o histograma apresenta caudas mais grossas, isto é, *fat tails* e o pico mais achatado do que a distribuição normal estebelece. Não pretendemos discutir neste estudo os motivos que causam este fenômeno. Mencionaremos apenas que para certos ativos esta não-normalidade tem raízes claras. Por exemplo, as taxas de câmbio podem apresentar distribuição de probabilidades muito diferentes da distribuição normal, como é o caso brasileiro, onde há muito tempo as taxas de câmbio estão restritas a flutuações segundo uma banda cambial estabelecida pelo Banco Central. Nos casos em que a distribuição de retornos não são normais, podem ocorrer problemas na utilização da fórmula de Black e Scholes para precificar opções, tendo em vista que uma hipótese fundamental do modelo foi violada.

O objetivo principal deste trabalho é sugerir e estudar um processo que modele a dinâmica do preço de ação, que, por um lado, seja coerente com os dados amostrais dos retornos do ativo-objeto e portanto, não necessite da suposição de normalidade dos retornos e que, por outro lado, permita a aplicação do argumento de Black e Scholes referente à construção de uma carteira livre de risco para a precificação de opções.

São propostos três modelos que serão discutidos em detalhes no Capítulo 4. O primeiro modelo chama-se *processo de mistura por variável* ou simplesmente *processo de mistura*. Neste processo, o

[1] veja Bibby e Sorensen

preço do ativo-objeto é dado por

$$S_t = S_0 \exp\{\mu_k t + \sigma_k W_t\}, t \geq 0, \tag{1.3}$$

com a probabilidade p_k.

Temos, portanto neste modelo, que k pode assumir valores $1, 2, \ldots, m$, e que $m, \mu_1, \ldots, \mu_m, \sigma_1, \sigma_m, p_1, \ldots, p_m$ são constantes fixas, com a restrição $p_1 + \ldots + p_m = 1$. O processo definido em (1.3) pode ser descrito da seguinte maneira: no momento $t = 0$ um dos m processos brownianos $\mu_k t + \sigma_k W_t, t \geq 0$, pode ser escolhido com a respectiva probabilidade p_k. O preço do ativo-objeto então seguirá o movimento browniano geométrico escolhido até o vencimento do contrato da opção.

Outro modelo a ser estudado chama-se *processo de mistura por processo pontual de Poisson*. Neste modelo $S.$ é um movimento browniamo geométrico em cada intervalo $[t_n, t_{n+1}]$, onde os instantes $t_0 = 0, t_1, t_2, \ldots$ são pontos do processo pontual de Poisson, sendo que em cada t_n a escolha do browniano que será seguido é feita de maneira independente e com mesma distribuição.

O terceiro modelo é parecido com o segundo, com a diferença de que os pontos $\{t_n, n \in \mathbb{N}\}$ são determinados por um processo pontual que aproxima o processo pontual de Poisson. Este modelo será denotado por *processo de mistura por processo pontual de Poisson aproximado.*

Observe que a função de densidade da distribuição dos retornos do processo de mistura é

$$\sum_{k=1}^{m} p_k f_k(x), \text{ onde } f_k(x) \text{ representa a densidade de } \mathcal{N}(\mu_k, \sigma_k^2) \tag{1.4}$$

Usando o fato de que uma mistura de densidades normais pode aproximar suficientemente bem qualquer densidade razoavelmente suave, concluímos que os parâmetros do processo de mistura podem ser escolhidos de forma que seus retornos tenham uma distribuição semelhante à distribuição empírica. A mesma conclusão vale para o processo de mistura por processo pontual de Poisson e para o processo de mistura por processo pontual de Possion aproximado, através da utilização de um argumento parecido com o aplicado para o processo de mistura simples. A questão da determinação dos valores de parâmetros da mistura será abordada no Capítulo 7.

Observe que todos os três modelos são compostos de movimentos brownianos. Isto foi estabelecido com a intenção de possibilitar a aplicação do argumento de Black e Scholes para estes modelos. No Capítulo 5, realizamos uma análise para verificar se realmente o argumento de Black e Scholes pode ser aplicado a estes modelos de comportamento de preço do ativo-objeto. Na Seção 5.1 são mostrados os detalhes sobre o argumento original. Na Seção 5.2 o argumento é aplicado para um modelo de dinâmica do preço do ativo-objeto que tem trajetórias diferenciáveis, visando enfatizar a importância e a necessidade dos conceitos de cálculo estocástico. A Seção 5.3 mostra que uma condição necessária para o funcionamento do argumento consiste no fato de que uma certa equação, que é um fruto da aplicação do argumento, deve ser uma equação diferencial determinística e não equação diferencial estocástica. Mostramos que isto implica a não-aplicabilidade do argumento de Black e Scholes para o caso em que pelo menos dois dos $\sigma_1, \ldots, \sigma_m$ são distintos. Para o caso quando todos eles são iguais, o argumento realmente funciona, exigindo-se porém certos cuidados técnicos. Para tratar deles, apresentamos no Capitulo 3 as ferramentas necessárias do cálculo estocástico. No Capítulo 6, serão abordados os principais passos necessários para a resolução da equação diferencial determinística que é resultado da aplicação de cálculo estocástico ao argumento de Black e Scholes.

Observação:

A presente versão do trabalho não se constitui na versão final e portanto é passível de correções. Em particular, na versão final, estarão presentes o Apêndice que conterá definições e resultados secundários que serão utilizados em nosso trabalho e a Bibliografia complementar que corresponderá aos estudos e publicações que foram pesquisados durante a elaboração desta dissertação. A melhor formatação do texto, com a sequência contínua das páginas e outras alterações de lay-out, também farão parte da versão final. Cabe ressaltar, porém, que os principais resultados do trabalho estão apresentados e discutidos nesta versão, representando de forma completa o escopo da dissertação.

Chapter 2

Conceitos financeiros

Neste capítulo, apresentaremos os principais conceitos financeiros que serão necessários para o estudo desenvolvido neste trabalho. Será realizada uma breve introdução contendo as características relevantes dos instrumentos financeiros a serem analisados: os contratos de opção. A primeira seção é dedicada à descrição destes contratos que constituem produtos flexíveis e extremamente úteis para a gestão financeira. Serão estudados os tipos de opção e sua avaliação na data de vencimento. Na segunda seção, apresentaremos o princípio da arbitragem, que representa um fundamento importante para o desenvolvimento da teoria de finanças e para a avaliação de derivativos. Aplicações do princípio de arbitragem serão mostradas a título de exemplificação. Finalmente, na terceira seção, será discutida a avaliação de derivativos segundo a hipótese teórica de neutralidade ao risco, onde supõe-se que todos os ativos têm o mesmo retono esperado. A neutralidade a risco representará uma forma alternativa ao raciocínio de Black e Scholes que permitirá o cálculo do preço de uma opção de compra.

2.1 Derivativos financeiros

Derivativos financeiros, ou simplesmente derivativos, são contratos financeiros cujo valor depende ou deriva de um ativo denominado ativo-objeto. O ativo-objeto pode ser qualquer produto, como por exemplo, commodities, automóveis ou qualquer instrumento financeiro, como taxa de juros, taxa de câmbio, ações de empresas, operações de crédito etc. Uma definição mais precisa de derivativo é dada abaixo (Ingersoll, 1987).

Definição 2.1 *Um contrato financeiro é um contrato derivativo se seu valor numa data de vencimento T puder ser determinado de maneira exata através do preço de mercado do ativo-objeto no instante T.*

Trataremos neste estudo da avaliação dos contratos de opções que são derivativos que conferem ao seu titular a alternativa de escolha entre exercer ou não exercer determinado direito. A principal característica das opções é permitir que o titular tome determinada atitude somente se lhe for conveniente. Assim, o titular da opção possui um direito sem uma obrigação correspondente. Este produto difere, portanto, dos demais produtos financeiros onde a um direito está sempre associado um dever correspondente.

A flexibilidade da opção a torna um produto único, permitindo que sejam realizadas diversas estratégias nos mercados financeiros cuja estruturação não seria possível com a utilização dos ativos ou outros derivativos convencionais, como por exemplo, futuros e swaps.

As inovações no segmento de derivativos estão basicamente atreladas ao desenvolvimento de produtos com características dos contratos de opções. Assim, os chamados derivativos exóticos que correspondem à onda mais recente de derivativos são, em sua maioria, produtos financeiros tradicionais com diversas cláusulas de opções embutidas.

Neste trabalho, buscaremos apresentar os aspectos mais relevantes para a precificação, ou seja, a avaliação do preço das opções, tendo em vista que um direito sem uma obrigação deve ter um valor de negociação.

Vamos concentrar nossos esforços na análise das opções tradicionais mais simples: as opções de compra européias. Com menor ênfase, serão tratadas as opções de venda, cuja precificação será decorrente da avaliação das opções de compra. O ativo-objeto, isto é, o ativo ao qual o derivativo se refere, será na maioria das vezes, para simplificar a exposição de conceitos sem a perda de generalidade, ações de companhias de capital aberto. A extensão dos resultados para outros ativos, desde que observadas suas características peculiares, pode ser realizada sem grandes problemas adicionais, principalmente em se tratando de taxa de câmbio ou commodities.

Vamos primeiramente estabelecer os conceitos relativos às opções, apresentando os principais critérios de classificação destes derivativos.

Em função do período de exercício do direito, podemos diferenciar dois tipos de opções: as opções do tipo europeu e as opções do tipo americano. As opções americanas conferem ao titular o direito de exercício a qualquer momento até uma data específica, chamada de data de vencimento da opção. As opções européias conferem o direito de exercício somente na data de vencimento.

Em função do tipo de direito, existem duas classificações genéricas: as opções de compra e as opções de venda. Uma opção de compra confere ao titular o direito de comprar, no futuro, uma determinada quantia de ações de uma companhia por um preço pré-estabelecido, chamado preço de exercício. A opção de compra também é comumente denominada de call. O titular da call européia portanto só exercerá seu direito se, na data do vencimento, o preço da ação estiver valendo mais do que o preço de exercício, uma vez que poderia estar adquirindo a ação por um preço menor que o negociado no mercado. Caso contrário, se o preço da ação no vencimento for menor que o preço de exercício, o titular da call não teria motivação para exercer seu direito de compra. Neste caso, o direito de compra não é exercido e diz-se que a opção virá pó.

De modo semelhante, uma opção de venda européia confere ao titular o direito de vender, na data de vencimento, uma determinada quantidade de ações de uma companhia por um preço pré-estabelecido, chamado preço de exercício. A opção de venda também recebe a denominação de put. O titular da put só exerce seu direito de venda se, na data de vencimento, o preço da ação estiver valendo menos que o preço de exercício, uma vez que possui o direito de vender as ações por um valor maior do que o negociado no mercado. Caso o preço de exercício seja menor do que o preço de mercado, então não há exercício e a opção vira pó.

O detentor de uma opção, de compra ou de venda, é denominado de titular, sendo o proprietário do direito de compra ou venda, respectivamente. A contraparte do titular é o lançador, ou seja, o indivíduo que vende o direito ao titular. Observe que um direito sem obrigação é valioso e portanto, o titular da opção deve pagar um valor, chamado de prêmio, para adquirir a call ou a put. O prêmio é pago ao lançador, pois este, ao negociar a opção, fica com uma obrigação sem direito. No caso de uma call, por exemplo, o lançador fica obrigado a vender as ações de uma companhia pelo preço de exercício se o titular exercer seu direito de compra. Obviamente, uma vez que o exercício da opção só será racional quando for desvantajoso ao lançador, é lógico que este deva cobrar um prêmio para assumir esta posição de obrigação sem direito.

Pelas características expostas acima, podemos estabelecer definições mais formais dos derivativos que serão tratados neste trabalho.

Definição 2.2 *Uma opção de compra do tipo europeu sobre um ativo-objeto de preço S_t é o direito de compra do ativo-objeto por um preço pré-estabelecido K chamado preço de exercício. Este direito só pode ser exercido numa data predeterminada T chamada de vencimento da opção. A opção de compra ou direito de compra numa data futura pode ser adquirida por um valor C_t chamado de prêmio.*

Definição 2.3 *Uma opção de venda do tipo europeu sobre um ativo-objeto de preço S_t é o direito de venda do ativo-objeto por um preço pré-estabelecido K chamado preço de exercício. Este direito só pode ser exercido numa data predeterminada T chamada de vencimento da opção. A opção de venda ou direito de venda numa data futura pode ser adquirida por um valor P_t chamado de prêmio.*

De acordo com a nomenclatura americana, conforme já mencionado anteriormente, uma opção de compra também pode ser chamada de call e uma opção de venda pode ser denominada put.

A partir das definições dadas acima, vamos agora estabelecer os valores das opções na data de vencimento. Estaremos utilizando o conceito de arbitragem para definirmos as condições , no vencimento, dos prêmios das opções. A arbitragem será o assunto de nosso estudo na próxima seção. Os resultados a seguir porém serão mostrados de forma simples, intuitiva, sem a necessidade de definição formal do princípio de arbitragem.

Supondo $t = 0$ o momento atual e $t = T$ a data de vencimento, com $T > 0$, então para uma opção de compra, temos:

$$C_T = \max(S_T - K, 0) \tag{2.1}$$

onde
C_T é o preço da opção de compra no vencimento
S_T é o preço do ativo-objeto no vencimento
K é o preço de exercício da opção de compra

De modo análogo, no caso de uma opção de venda, no vencimento o prêmio é dado por:

$$P_T = \max(K - S_T, 0) \tag{2.2}$$

onde
P_T é o preço da opção de venda no vencimento
K é o preço de exercício da opção de venda

A relação acima (2.1) representa o prêmio avaliado na data T, ou seja, o valor a ser pago em T para a aquisição de um direito de comprar o ativo-objeto por um preço K na própria data T de vencimento da opção. Podemos mostrar que esta relação é válida, de modo informal através de um raciocínio simples, explicado a seguir.

Suponha que na data de vencimento T, $C_T > \max(S_T - K, 0)$. Exemplificando com dados numéricos, sejam o preço do ativo-objeto $S_T = 120$, o prêmio da opção de compra do ativo-objeto $C_T = 30$ e o preço de exercício $K = 100$. Nestas condições de mercado, como $C_T \neq \max(S_T - K, 0)$ então há oportunidades de arbitragem, isto é, ganhos sem risco. Se um agente vender a opção de compra, obterá um fluxo de caixa positivo de 30. Sendo lançador de opção, este agente ficará na obrigação de atender o titular no caso deste escolher pelo exercício de seu direito. Como $S_T = 120 > 100 = K$, é racional esperar que o titular exerça seu direito de compra do ativo-objto por 100, uma vez que poderia imediatamente vendê-lo pelo preço de mercado, maior, de 120. Com o exercício da opção, o lançador terá a obrigação de vender o ativo-objeto por 100, recebendo portanto um fluxo de caixa positivo de 100. Para isto, terá de adquirir o ativo no mercado à vista, pagando o preço de mercado e portanto, desembolsando 120. O fluxo líquido do lançador será positivo e igual a $30 + 100 - 120 = 10$. O titular, caso se desfizesse imediatamente das ações que adquiriu pelo exercício da opção, teria um fluxo líquido de $-30 - 100 + 120 = -10$. Uma vez que no vencimento T todos os parâmetros S_T, K, C_T são conhecidos, não existe incerteza associada à estratégia do lançador. Portanto, a estratégia de venda da opção por 30, compra da ação por 120 e venda da ação por 100 estaria lhe garantindo um lucro de 10 sem risco. Em finanças, suporemos que este tipo de oportunidade de ganho sem risco não possa existir. Intuitivamente, parece ser claro que um possível comprador da call não estaria disposto a pagar, no vencimento, 30 pelo direito de comprar um ativo por 100, sabendo que o mercado o está negociando por 120. Assim,

$$C_T \not> \max(S_T - K, 0) \tag{2.3}$$

Um raciocínio análogo, pode ser realizado se supusermos uma situação oposta, onde $C_T < \max(S_T - K, 0)$, por exemplo, $S_T = 120$, $C_T = 10$ e $K = 100$. Assim, a compra de uma call e seu exercício imediato, pois $t = T$, conferiria ao agente um fluxo de caixa líquido positivo sem risco. Pela compra da opção desembolsaria 10, pelo exercício desembolsaria mais 100 para comprar o ativo-objeto e pela venda imediata deste ativo no mercado, receberia 120. Em termos líquidos, teria $-10 - 100 + 120 = 10$, obtendo um ganho sem risco. Adicionalmente, o prêmio da opção não pode assumir valores negativos, tendo em vista que não é racional adquirir um direito que implique em perdas. Com isso

$$C_T \not< \max(S_T - K, 0) \tag{2.4}$$

Das equações (2.3) e (2.4), $C_T = \max(S_T - K, 0)$ e portanto a equação (2.1) vale. Através de um raciocínio semelhante, podemos mostrar que a equação (2.2) também é válida.

Observe que no vencimento, os valores da call e da put são facilmente identificáveis. O principal problema na avaliação de derivativos é a obtenção de seus valores em momentos anteriores ao vencimento. Por existir uma incerteza com relação ao preço que o ativo-objeto terá no vencimento, torna-se necessária a utilização da teoria de probabilidade para avaliarmos o preço da opção.

Obviamente, o preço da opção dependerá do modelo probabilístico adotado para descrever a distribuição do preço do ativo-objeto e da dinâmica da alteração de preço no decorrer do tempo. Na maioria dos estudos, inclusive no artigo de Black e Scholes, supõe-se que, dado $t \geq 0$:

$$S_t = S_0 \exp\left[\left(\mu - \frac{\sigma}{2}\right)t + \sigma W_t\right] \tag{2.5}$$

onde

t representa o tempo

S_t representa o preço do ativo-objeto no tempo t

μ é uma constante que caracteriza a taxa instantânea de retorno do ativo-objeto

σ é uma outra constante que representa o grau de dispersão dos retornos do ativo-objeto, também chamada de volatilidade.

W_t, $t \geq 0$, é um processo estocástico que se chama processo de Wiener ou movimento browniano

A equação (2.5) pode ser escrita em termos de uma equação diferencial estocástica dada por:

$$\frac{dS}{S} = \mu dt + \sigma dW_t \tag{2.6}$$

com $t \geq 0$.

A partir da equação (2.5) segue, em particular, que na data de vencimento da opção $t = T$, considerando uma data de referência $t = 0$, S_T tem distribuição log-normal com média $\mu - \frac{\sigma}{2}$ e variância σ^2. Observaremos que é possível precificar a opção de compra sabendo-se a distribuição de S_T sem a necessidade de imposição de qualquer hipótese sobre a dinâmica do preço do ativo-objeto. Tal forma de precificação utiliza o princípio de mundo *risco neutro* que será apresentado na Seção 2.3.

A equação (2.6), que denota uma diferencial estocástica, será analisada posteriormente de forma mais detalhada. Em linhas gerais, a equação indica que o retorno do ativo-objeto é resultado de dois fatores: uma taxa de retorno constante representando uma tendência de valorização do ativo e um componente aleatório caracterizando o comportamento de risco do ativo. A precificação de opções de compra para o modelo proposto pela equação (2.5) foi realizada por Black e Scholes e será descrita e discutida neste trabalho no Capítulo 5. Com relação ao modelo do preço do ativo-objeto, convém destacar que existem evidências empíricas de que S_T desvia-se significativamente de uma distribuição log-normal. Em outras palavras, (2.5) e consequentemente (2.6) podem não configurar processos apropriados para o modelamento do preço do ativo-objeto.

Assim, neste trabalho avançaremos também o estudo de Black e Scholes apresentando uma sugestão de um outro modelo. Este modelo permitirá a derivação de uma fórmula de precificação de opções para uma suposição de que o preço do ativo-objeto segue um processo de difusão baseado numa mistura de brownianos. A formulação teórica para este novo modelo será desenvolvida nos Capítulos 4 e 5, o algoritmo de implementação será apresentado no Capítulo 7.

2.2 Princípio de arbitragem

Uma vez apresentados os derivativos que serão foco de nosso estudo, vamos definir a representação dos elementos financeiros fundamentais para o entendimento do conceito de arbitragem. A arbitragem constitui um importante conceito em finanças, com implicações extremamente poderosas para a avaliação de instrumentos financeiros.

Consideremos A_t o vetor que representa os preços de todos os ativos financeiros negociados no mercado em um instante t. Assim,

$$A_t = \begin{pmatrix} A_1(t) \\ \vdots \\ A_N(t) \end{pmatrix} \tag{2.7}$$

onde

N é o número total de ativos

$A_1(t)$ pode representar o preço de um instrumento de renda fixa livre de risco

$A_2(t)$ pode representar o preço de uma ação

$A_3(t)$ o preço de uma opção sobre ações e assim por diante.

Os preços dos ativos estão sujeitos a alterações em decorrência do comportamento dos mercados. Assim, por exemplo, novas informações sobre a economia nacional ou internacional, movimentações políticas ou alterações específicas de setores ou empresas, implicam mudanças de expectativas ou percepções de risco e retorno, podendo causar variações de preços nos ativos.

Vamos definir V como o vetor que representa todos os possíveis estados da natureza ou situações que podem vir a ocorrer em um determinado instante no futuro. Assim, considerando um número limitado de estados da natureza:

$$V = \begin{pmatrix} \omega_1 \\ \vdots \\ \omega_k \end{pmatrix} \tag{2.8}$$

onde cada ω_i, $1 \leq i \leq k$, representa um evento ou estado da natureza que pode ocorrer.

Uma observação importante é que os estados da natureza são mutuamente exclusivos e conjuntamente excludentes, cobrindo portanto, todos os possíveis cenários futuros. Sob cada cenário, os ativos terão um preço distinto, uma vez que as expectativas de risco e retorno são dependentes do estado da natureza. Estabeleçamos D como a matriz que representa o resultados de cada ativo sob cada possível estado da natureza no momento t^*. Temos:

$$D = \begin{pmatrix} A_{11}(t^*) & \cdots & A_{1k}(t^*) \\ \vdots & \vdots & \vdots \\ A_{N1}(t^*) & \cdots & A_{Nk}(t^*) \end{pmatrix} \tag{2.9}$$

onde $A_{ij}(t^*)$, $1 \leq i \leq N$ e $1 \leq j \leq k$, é o resultado do ativo i no estado da natureza j.

Portanto, as linhas da matriz D correspondem aos resultados de um ativo i em cada um dos k possíveis estados da natureza, enquanto as colunas correspondem aos resultados de cada um dos N ativos para um determinado estado da natureza j. Em finanças, os resultados A_{ij} dos ativos possuem 2 componentes. O primeiro componente do resultado refere-se ao ganho ou à perda de capital decorrente da valorização ou desvalorização da posição mantida no ativo. O segundo componente refere-se ao rendimento corrente que corresponde ao valor recebido pela manutenção de uma posição no ativo, como por exemplo, dividendos e juros. Podemos, simplificadamente, estabelecer que cada A_{ij} corresponderá ao valor resultante da alteração do preço do ativo i sob o estado da natureza j ajustado por todos os rendimentos distribuídos ou pagos pelo ativo.

Até o momento, apresentamos as notações referentes aos preços dos ativos e os possíveis estados da natureza. Os agentes financeiros, em geral, aplicam ou captam recursos utilizando-se dos ativos disponíveis para negociação. Uma carteira de investimento Θ pode ser definida como uma combinação de ativos representada por dois parâmetros: a posição estratégica em cada ativo e a quantidade da exposição ao ativo. Temos então:

$$\Theta = \begin{pmatrix} \Theta_1 \\ \vdots \\ \Theta_N \end{pmatrix} \tag{2.10}$$

onde o sinal de Θ_i, $1 \leq i \leq N$, representa o posicionamento estratégico, isto é, uma posição comprada (sinal positivo) ou vendida (sinal negativo) no ativo i e o valor absoluto de Θ_i corresponde à quantidade do ativo i na carteira. Obviamente, se $\Theta_i = 0$, então o ativo i não está presente na carteira.

De acordo com as definições anteriores, o valor $P(t)$ de uma carteira no instante t é dado por

$$P(t) = A_t^T \Theta = \sum_{i=1}^{N} A_i(t)\Theta_i \tag{2.11}$$

onde A_t^T é a transposta de A_t, ou seja, o valor da carteira é simplesmente a soma dos valores dos ativos ajustados pelas posições Θ_i.

Adicionalmente, podemos definir o vetor R_{t^*} como representante dos possíveis resultados da carteira Θ para cada um dos k estados da natureza, no instante t^*.

$$R_{t^*} = D^T\Theta = \begin{pmatrix} A_{11}(t^*) & \cdots & A_{N1}(t^*) \\ \vdots & \vdots & \vdots \\ A_{1k}(t^*) & \cdots & A_{Nk}(t^*) \end{pmatrix} \begin{pmatrix} \Theta_1 \\ \vdots \\ \Theta_N \end{pmatrix} = \begin{pmatrix} R_1(t^*) \\ \vdots \\ R_k(t^*) \end{pmatrix} \tag{2.12}$$

onde R_k é o resultado da carteira no estado da natureza k, com $R_k = \sum_{i=1}^{N} A_{ik}\Theta_i$.

Com base nos conceitos expostos anteriormente e considerando um modelo simplificado de dois períodos, onde t é o instante inicial e t^*, o instante final, vamos definir a arbitragem.

Definição 2.4 *Uma estratégia de arbitragem consiste na estruturação de uma carteira tal que uma das seguintes condições seja satisfeita:*

$$A^T\Theta \leq 0$$

e

$$D^T\Theta > 0 \tag{2.13}$$

ou

$$A^T\Theta < 0$$

e

$$D^T\Theta \geq 0 \tag{2.14}$$

onde as desigualdades em (2.13) e (2.14) são vetoriais.

Ou seja, a arbitragem é uma estratégia que permite que uma carteira de valor inicial nulo ou negativo gere resultados estritamente positivos, independentemente do estado da natureza. Assim, com um investimento inicial de custo nulo ou negativo (recebimento de fluxo de caixa), a arbitragem representa uma estratégia que confere, em um momento posterior, fluxos de caixa estritamente maiores que zero. De modo análogo, a arbitragem permite que dado um recebimento inicial de fluxo de caixa, num momento posterior, há um recebimento de fluxo de caixa nulo ou positivo. A arbitragem portanto representa a obtenção de ganhos, na forma de entradas líquidas de fluxos de caixa ajustados pelo valor do dinheiro no tempo, sem risco. Em um mercado em equilíbrio é racional supor que não haja possibilidade de arbitragem, isto é, os preços de todos os ativos possuem valores tais que não permitam que os agentes financeiros possam obter ganhos sem risco. Obviamente, este resultado parece ser lógico, uma vez que um possível desequilíbrio de preços causaria uma movimentação dos agentes financeiros para aproveitar oportunidades de arbitragem, fazendo com que os ativos, pela pressão de demanda e oferta, fossem conduzidos a preços de equilíbrio.

O resultado abaixo corresponde ao princípio de arbitragem que será de extrema importância para o desenvolvimento deste trabalho.

Princípio de arbitragem *Num mercado em equilíbrio, não existem possibilidades de arbitragem*

Este resultado é decorrente da definição dos conceitos expostos anteriormente e da definição de equilíbrio de mercado.

A seguir, apresentaremos dois teoremas que explicitam matematicamente o conceito de arbitragem. Maiores detalhes sobre a obtenção e validade do princípio de arbitragem e de suas origens e consequências podem ser encontrados em Dothan (Dothan, 1996).

Teorema 2.1 *Se não há oportunidades de arbitragem, então existe $\psi > 0$ tal que $A = D\psi$.*

Teorema 2.2 *Se existe $\psi > 0$ tal que $S = D\psi$, então não existem oportunidades de arbitragem.*

2.2.1 Exemplo de precificação de derivativo

O problema de precificação de derivativos consiste na determinação de seu preço num instante $t < T$, onde T representa a data de vencimento do derivativo.

Vamos utilizar o teorema 2.2 para exemplificarmos o conceito de arbitragem. Neste exemplo, precificaremos uma opção de compra no instante t sob condições simplificadas, onde o número de estados da natureza possíveis é limitado e o vencimento do derivativo ocorre em um instante posterior t^*. Vamos analisar somente o comportamento dos ativos nos dois períodos t e t^*.

Consideremos a seguinte representação de 3 ativos e 2 possíveis estados futuros da natureza. Supondo que não existam oportunidades de arbitragem, então de acordo com o teorema 2.2:

$$A_t = D\psi \tag{2.15}$$

Supondo

$$D = \begin{pmatrix} A_1(t) \\ A_2(t) \\ A_3(t) \end{pmatrix} = \begin{pmatrix} A_{11}(t^*) & A_{12}(t^*) \\ A_{21}(t^*) & A_{22}(t^*) \\ A_{31}(t^*) & A_{32}(t^*) \end{pmatrix}$$

e

$$\psi = \begin{pmatrix} \psi_1 \\ \psi_2 \end{pmatrix}$$

então:

$$A_t = \begin{pmatrix} A_1(t) \\ A_2(t) \\ A_3(t) \end{pmatrix} = \begin{pmatrix} A_{11}(t^*) & A_{12}(t^*) \\ A_{21}(t^*) & A_{22}(t^*) \\ A_{31}(t^*) & A_{32}(t^*) \end{pmatrix} \begin{pmatrix} \psi_1 \\ \psi_2 \end{pmatrix} \tag{2.16}$$

onde

$A_1(\cdot)$ é o preço de um instrumento de renda fixa cuja rentabilidade é a taxa e juros livre de risco r

$A_2(\cdot)$ é o preço de uma ação de uma determinada empresa
$A_3(\cdot)$ é o preço de uma opção sobre esta ação

Estamos, sem perda de generalidade, supondo que a taxa de juros livre de risco é capitalizada continuamente, ou seja, a cada instante os juros são incorporados ao valor do ativo livre de risco.

Para o instrumento de renda fixa livre de risco, independente do estado da natureza futuro, temos:

$$A_{11}(t^*) = A_{12}(t^*) = A_1(t)\exp[r(t^* - t)] \tag{2.17}$$

Assim,

$$\begin{pmatrix} A_1(t) \\ A_2(t) \\ A_3(t) \end{pmatrix} = \begin{pmatrix} A_1(t)\exp[r(t^* - t)] & A_1(t)\exp[r(t^* - t)] \\ A_{21}(t^*) & A_{22}(t^*) \\ A_{31}(t^*) & A_{32}(t^*) \end{pmatrix} \begin{pmatrix} \psi_1 \\ \psi_2 \end{pmatrix} \tag{2.18}$$

Podemos simplificar a equação acima eliminando $A_1(t)$:

$$\begin{pmatrix} 1 \\ A_2(t) \\ A_3(t) \end{pmatrix} \begin{pmatrix} \exp[r(t^* - t)] & \exp[r(t^* - t)] \\ A_{21}(t^*) & A_{22}(t^*) \\ A_{31}(t^*) & A_{32}(t^*) \end{pmatrix} \begin{pmatrix} \psi_1 \\ \psi_2 \end{pmatrix} \tag{2.19}$$

Vamos introduzir alguns dados numéricos para mostrar os procedimentos para precificação da opção de compra. Suponhamos que o mercado esteja negociando a ação, no instante t, por 100, ou seja: $A_2(t) = 100$. Consideremos assim que, dependendo do cenário futuro em t^*, o resultado da ação seja conhecido, podendo ser 80 ou 130. Por exemplo, $A_{21}(t^*) = 80$ ou $A_{22}(t^*) = 130$. Ainda, vamos supor que a taxa de juros livre de risco é 10% ao período referente à escala de medição de t.

Se a opção de compra sobre a ação vencer em $t^* = T$, então, de acordo com os resultados apresentados na seção anterior, seu valor será:

$$A_{3\cdot} = \max(A_{2\cdot}(t^*) - K, 0) \tag{2.20}$$

Vamos supor que o preço de exercício seja 100, isto é, $K = 100$. Assim, temos, de acordo com a condição no vencimento para a opção de compra dada pela equação (2.1):

$$A_{31}(t^*) = \max(A_{21}(t^*) - K, 0) = \max(80 - 100, 0) = 0 \tag{2.21}$$

e

$$A_{32}(t^*) = \max(A_{22}(t^*) - K, 0) = \max(1300 - 100, 0) = 30 \tag{2.22}$$

Para fins de exposição, sem prejuízo da generalidade, uma vez que poderíamos ajustar a taxa livre de risco por uma escala apropriada, consideremos $t^* - t = 1$. Portanto, substituindo as variáveis pelos valores do exemplo, temos:

$$\begin{pmatrix} 1 \\ 100 \\ A_3(t) \end{pmatrix} = \begin{pmatrix} \exp(0,1) & \exp(0,1) \\ 80 & 130 \\ 0 & 30 \end{pmatrix} \begin{pmatrix} \psi_1 \\ \psi_2 \end{pmatrix} \tag{2.23}$$

Observe que o prêmio da call no instante t, $A_3(t)$, não foi dado. Em uma situação em que não exista possibilidade de arbitragem, $A_3(t)$ pode ser calculado a partir dos preços dos outros ativos negociados no mercado. Passemos então ao cálculo de $A_3(t)$ que representa o prêmio da opção de compra.

A representação acima corresponde ao seguinte sistema de equações.

$$\begin{cases} 1 = \exp(0,1)\psi_1 + \exp(0,1)\psi_2 \\ 100 = 80\psi_1 + 130\psi_2 \\ A_3(t) = 0 \times \psi_1 + 30 \times \psi_2 \end{cases} \tag{2.24}$$

Resolvendo o sistema acima, podemos a partir das duas primeiras equações, calcular ψ_1 e ψ_2. Assim,

$\psi_1 = 0,352577$ e $\psi_2 = 0,552260$

Como $\psi_1 > 0$ e $\psi_2 > 0$, pelo teorema 2.2, então não há, de acordo com os preços vigentes e resultados futuros para cada cenário, oportunidades de arbitragem.

A partir destes valores, vamos calcular o preço $A_3(t)$ da opção de compra no instante t. Substituindo ψ_1 e ψ_2 na terceira equação do sistema (2.24),

$$A_3(t) = 0 \times 0,352577 + 30 \times 0,552260 \approx 16,5678 \tag{2.25}$$

O valor $A_3(t) = 16,57$, arredondando para suprimir frações menores que centavos, representa portanto, o preço pelo qual teoricamente a call deveria ser negociada no instante t de modo a não existirem oportunidades de arbitragem. Podemos verificar esta afirmação, supondo $A_3(t) \neq 16,57$. Inicialmente, consideremos que o mercado esteja em desequilíbrio, negociando a opção no instante t por $A_3(t) = 15,00 < 16,57$. Nesta situação, o mercado está sub-avaliando a opção. Um arbitrador poderia estruturar a seguinte estratégia, supondo que partes não inteiras de ativos possam ser negociadas: comprar $1,6$ opção, vender $1,0$ ação e aplicar $76,0$ em um ativo de renda fixa livre de risco, ou seja:

$$\Theta = \begin{pmatrix} 76,0 \\ -1,0 \\ 1,6 \end{pmatrix} \tag{2.26}$$

Com isso, o valor do fluxo de caixa inicial para a montagem desta carteira é $P(t) = A^T\Theta$, isto é:

$$P(t) = A^T\Theta = \begin{pmatrix} 1 & 100 & 15 \end{pmatrix} \begin{pmatrix} 76,0 \\ -1,0 \\ 1,6 \end{pmatrix} = 1 \times 76,0 + 100 \times (-1,0) + 15 \times (1,6) = 0 \tag{2.27}$$

No instante t^*, dois possíveis estados da natureza podem ocorrer. No primeiro cenário, temos $A_{11} = \exp(0,1)$, $A_{21} = 80$ e $A_{31} = 0$. No segundo cenário, temos $A_{12} = \exp(0,1)$, $A_{22} = 130$ e $A_{32} = 30$. Vamos avaliar o resultado propiciado pela estratégia Θ em t^*. Temos em t^*:

$$R_{t^*} = D^T\Theta = \begin{pmatrix} \exp(0,1) & 80 & 0 \\ \exp(0,1) & 130 & 30 \end{pmatrix} \begin{pmatrix} 76,0 \\ -1,0 \\ 1,6 \end{pmatrix} \tag{2.28}$$

que equivale a:

$$R_{t^*} = \begin{pmatrix} \exp(0,1) \times 76 + 80 \times (-1) + 0 \times 1,6 \\ \exp(0,1) \times 76,0 + 130 \times (-1,0) + 30 \times 1,6 \end{pmatrix} = \begin{pmatrix} 4,00 \\ 2,00 \end{pmatrix} > \begin{pmatrix} 0 \\ 0 \end{pmatrix} \tag{2.29}$$

Observe que $A^T\Theta = 0$ e $D^T\Theta > 0$ e portanto, segundo a definição (2.4), existem oportunidades de arbitragem. De fato, com a estratégia Θ, sendo $A_3(t) = 15,00 \neq 16,57$, pode-se a partir de um investimento inicial nulo, obter resultados positivos indiferente do estado da natureza em t^*: resultado $2,00$ no caso em que a ação passe a valer 130 e resultado $4,00$ quando a ação for para 80.

Consideremos agora o caso em que o mercado esteja negociando a opção por $A_3(t) = 20,00 > 16,57$, ou seja, a opção encontra-se super-avaliada. Nestas condições, também existem possibilidades de arbitragem, como por exemplo, a seguinte estratégia: venda de $2,0$ opções de compra, captação de $60,0$ através de um instrumento de renda fixa livre de risco e compra de $1,0$ ação. Esta estratégia implica um investimento inicial em t nulo e resultados estritamente positivos em t^*, independentemente do estado da natureza, pois sendo:

$$\Theta = \begin{pmatrix} -60,0 \\ 1,0 \\ -2,0 \end{pmatrix} \tag{2.30}$$

então:

$$P(t) = A^T\Theta = \begin{pmatrix} 1 & 100 & 20 \end{pmatrix} \begin{pmatrix} -60,0 \\ 1,0 \\ -2,0 \end{pmatrix} = -60,0 + 100,0 - 40,0 = 0 \tag{2.31}$$

e

$$R_{t^*} = D^T\Theta = \begin{pmatrix} \exp(0,1) & 80 & 0 \\ \exp(0,1) & 130 & 30 \end{pmatrix} \begin{pmatrix} -60,0 \\ 1,0 \\ -2,0 \end{pmatrix} \tag{2.32}$$

Ou seja,

$$R_{t^*} = \begin{pmatrix} \exp(0,1) \times (-60,0) \; 80 \times 1,0 \; 0 \times (-2,0) \\ \exp(0,1) \times (-60,0) \; 130 \times 1,0 \; 30 \times (-2,0) \end{pmatrix} = \begin{pmatrix} 13,69 \\ 3,69 \end{pmatrix} > \begin{pmatrix} 0 \\ 0 \end{pmatrix} \tag{2.33}$$

Novamente $A^T\Theta = 0$ e $D^T\Theta > 0$ e portanto, existe possibilidade de arbitragem. A partir de (2.29) e (2.33), o preço justo da opção, de acordo com as considerações anteriores pode ser calculado através dos preços dos outros ativos e para não haver oportunidades de arbitragem,

$$A_3(t) = 16,57 \tag{2.34}$$

Observação. ESte exemplo mostrou uma forma de precificação de opções de compra obedecendo-se o princípio de arbitragem. Ponto fundamental para a precificação neste caso foi a determinação de 2 possíevis valores do ativo-objeto nos dois únicos cenários distintos em $t*$. Os valores 80 e 130 para o ativo-objeto foram estabelecidos de forma arbitrária. Na Capítulo 4 veremos alguns processos que podem modelar o comportamento do preço do ativo-objeto que permitem a identificação de fórmulas de precificação de derivativos.

2.2.2 Put-Call parity

A opção de compra e a opção de venda têm prêmios que estão relacionados de acordo com a put-call parity. Esta relação entre os prêmios é decorrente do princípio de arbitragem. Nos capítulos posteriores, iremos calcular o valor de uma opção de compra. A opção de venda pode ser calculada a partir do preço da opção de compra utilizando-se a put-call parity, que será apresentada a seguir.

Vamos considerar os seguintes ativos: instrumento de renda fixa livre de risco, ação de uma empresa, opção de compra sobre esta ação e opção de venda sobre esta mesma ação. Temos então:

$$A_t = \begin{pmatrix} A_1(t) \\ A_2(t) \\ A_3(t) \\ A_4(t) \end{pmatrix} \tag{2.35}$$

onde, mudando a notação para melhor identificação da relação entre a opção de compra e a opção de venda:

$A_1(t) = 1$ representa o valor do ativo livre de risco

$A_2(t) = S$ representa o valor da ação

$A_3(t) = C$ é o prêmio da opção de compra, com data de vencimento em t^* e com preço de exercício K

$A_4(t) = P$ é o prêmio da opção de venda, com data de vencimento em t^* e com preço de exercício K

Observe que a suposição de que o ativo livre de risco vale 1 não restringe nossa análise, uma vez que podem ser criados títulos de renda fixa livre de risco múltiplos ou submúltiplos de $A_1(t)$. Assim, utilizando a nova notação:

$$A_t = \begin{pmatrix} 1 \\ S \\ C \\ P \end{pmatrix} \tag{2.36}$$

Vamos considerar que em t^*, na data de vencimento, existam k estados possíveis da natureza. Em cada estado, os possíveis resultados dos ativos serão dados pela matriz D, sendo r a taxa de juros livre de risco. Com isso, considerando um processo de somente dois períodos t e t^*:

$$D = \begin{pmatrix} \exp[r(t^* - t)] & \cdots & \exp[r(t^* - t)] \\ S_1(t^*) & \cdots & S_k(t^*) \\ C_1(t^*) & \cdots & C_k(t^*) \\ P_1(t^*) & \cdots & P_k(t^*) \end{pmatrix} \tag{2.37}$$

Estabeleceremos agora a seguinte estratégia:

$$\Theta = \begin{pmatrix} -K\exp[-r(t^* - t)] \\ 1 \\ -1 \\ 1 \end{pmatrix} \tag{2.38}$$

que representa uma aplicação em um ativo de renda fixa livre de risco equivalente a $K\exp[-r(t^*-t)]$, uma compra de 1 ação, uma venda de 1 opção de compra e uma compra de 1 opção de venda.

O investimento inicial é portanto:

$$P(t) = A^T\Theta = \begin{pmatrix} 1 & S & C & P \end{pmatrix} \begin{pmatrix} -K\exp[-r(t^*-t)] \\ 1 \\ -1 \\ 1 \end{pmatrix} = -K\exp[-r(t^*-t)] + S - C + P \tag{2.39}$$

O resultado dos ativos no momento t^* é:

$$D = \begin{pmatrix} \exp[r(t^*-t)] & \cdots & \exp[r(t^*-t)] \\ S_1(t^*) & \cdots & S_k(t^*) \\ C_1(t^*) & \cdots & C_k(t^*) \\ P_1(t^*) & \cdots & P_k(t^*) \end{pmatrix} \tag{2.40}$$

Dadas as equações (2.1) e (2.2), então:

$$D = \begin{pmatrix} \exp[r(t^*-t)] & \cdots & \exp[r(t^*-t)] \\ S_1(t^*) & \cdots & S_k(t^*) \\ \max(S_1(t^*)-K,0) & \cdots & \max(S_k(t^*)-K,0) \\ \max(K-S_1(t^*),0) & \cdots & \max(K-S_k(t^*),0) \end{pmatrix} \tag{2.41}$$

O resultado da estratégia Θ no vencimento das opções, t^*, é:

$$R_{t^*} = D^T\Theta = \begin{pmatrix} \exp[r(t^*-t)] & S_1 & C_1 & P_1 \\ vdots & \vdots & \vdots & \vdots \\ \exp[r(t^*-t)] & S_k & C_k & P_k \end{pmatrix} \begin{pmatrix} -K\exp[-r(t^*-t)] \\ 1 \\ -1 \\ 1 \end{pmatrix} \tag{2.42}$$

sendo

$$\begin{cases} C_1(t^*) = \max(S_1(t^*)-K,0) \\ \vdots \\ C_k(t^*) = \max(S_k(t^*)-K,0) \end{cases} \tag{2.43}$$

e

$$\begin{cases} P_1(t^*) = \max(K-S_1(t^*),0) \\ \vdots \\ P_k(t^*) = \max(K-S_k(t^*),0) \end{cases} \tag{2.44}$$

Quando a opção de compra não é exercida, então o preço da ação é menor que o preço de exercício. Neste caso, a opção de venda deve ser exercida. Analogamente, quando a opção de compra é exercida, não é racional exercer a opção de venda, pois o preço do ativo é necessariamente menor que o preço de exercício. Assim, considerando (2.43) e (2.44):

$$R_{t^*} = D^T\Theta = \begin{pmatrix} -K & S_1 & -C_1 & P_1 \\ \vdots & \vdots & \vdots & \vdots \\ -K & S_k & -C_k & P_k \end{pmatrix} = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix} \tag{2.45}$$

Para não haver possibilidade de arbitragem, pela definição (2.4), dado que $D^T\Theta = 0$, então $A^T\Theta = 0$. Portanto:

$$-K\exp[-r(t^* - t)] + S - C + P = 0 \tag{2.46}$$

Ou seja,

$$P = K\exp[-r(t^* - t)] - S + C \tag{2.47}$$

A equação acima representa a put-call parity, ou seja, a relação de paridade entre os prêmios da opção de compra e de venda, de forma a não haver oportunidades de arbitragem. Este resultado pode ser interpretado em termos financeiros. A estratégia de compra de uma opção de venda é equivalente a uma estratégia de aplicação de $K\exp[-r(t^* - t)]$ em um ativo de renda fixa livre de risco, venda de 1 ação e compra de 1 opção sobre a ação, com mesmo preço de exercício K da put. Em algumas ocasiões, podemos utilizar a equação (2.47) para sintetizar, ou seja, construir uma opção de venda a partir de outros ativos disponíveis para negociação. Este é o caso do mercado brasileiro no qual a opção de venda não possui liquidez e portanto é difícil de ser negociada, podendo prejudicar a administração de carteiras. Porém, através da put-call parity, este problema pode ser resolvido através da sintetização da opção de venda.

2.2.3 Carteira livre de risco

Nesta seção, descreveremos um conceito que é fundamental para o entendimento da derivação da fórmula de Black e Scholes. Faremos uma suposição de que é possível estabelecermos, através de quantidades balanceadas de ações e opções, uma carteira livre de risco. Entende-se por carteira livre de risco como um posicionamento Θ apropriado em ativos, cujo resultado seja independente do estado da natureza e portanto, não apresenta risco de variação do preço do ativo-objeto. Seja A o vetor que representa os preços do ativo de renda fixa livre de risco e da carteira livre de risco composta por ações e opções. Temos:

$$A_t = \begin{pmatrix} A_1(t) \\ A_2(t) \end{pmatrix} \tag{2.48}$$

onde
$A_1(t)$ é o preço do ativo de renda fixa livre de risco
$A_2(t)$ é o valor da carteira livre de risco

Observe que podemos considerar uma carteira de ativos como um ativo qualquer. Consideremos agora a seguinte estratégia:

$$\Theta = \begin{pmatrix} -\frac{A_2}{A_1} \\ 1 \end{pmatrix} \tag{2.49}$$

que corresponde a uma captação através da emissão de um título de renda fixa para a compra da carteira livre de risco. O investimento inicial para a estruturação desta estratégia é:

$$P(t) = A^T\Theta = \begin{pmatrix} A_1(t) & A_2(t) \end{pmatrix} \begin{pmatrix} -\frac{A_2}{A_1} \\ 1 \end{pmatrix} = -A_2 + A_2 = 0 \tag{2.50}$$

A matriz de resultados que representa o ativo livre de risco e a carteira livre de risco é, no momento t^*:

$$D = \begin{pmatrix} A_{11}(t^*) & \cdots & A_{1k}(t^*) \\ A_{21}(t^*) & \cdots & A_{2k}(t^*) \end{pmatrix} \tag{2.51}$$

Uma vez que o instrumento de renda fixa confere uma rentabilidade r temos: $A_{11}(t^*) = \cdots = A_{1k}(t^*) = A_1(t)\exp[r(t^* - t)]$. Como a carteira livre de risco conduz a resultados iguais, independentemente do estado da natureza, então é possível encontrar um único R tal que:

$$R = \frac{\ln \frac{A_{2\cdot}(t^*)}{A_2(t)}}{(t^* - t)} \tag{2.52}$$

Assim,

$$D = \begin{pmatrix} A_1\exp[r(t^* - t)] & \cdots & A_1\exp[r(t^* - t)] \\ A_2\exp[R(t^* - t)] & \cdots & A_2\exp[R(t^* - t)] \end{pmatrix} \tag{2.53}$$

O resultado da estratégia Θ é portanto, independente do cenário:

$$\begin{aligned} D^T\Theta &= \begin{pmatrix} A_1\exp[r(t^* - t)] & A_2\exp[R(t^* - t)] \\ \vdots & \vdots \\ A_1\exp[r(t^* - t)] & A_2\exp[R(t^* - t)] \end{pmatrix} \begin{pmatrix} -\frac{A_2}{A_1} \\ 1 \end{pmatrix} \\ &= \begin{pmatrix} -A_2\exp[r(t^* - t)] + A_2\exp[R(t^* - t)] \\ \vdots \\ -A_2\exp[r(t^* - t)] + A_2\exp[R(t^* - t)] \end{pmatrix} \end{aligned} \tag{2.54}$$

Como para os k estados da natureza, o resultado da estratégia é sempre o mesmo, para não haver possibilidade de arbitragem, uma vez que $A^T\Theta \leq 0$, então $D^T\Theta = 0$. Assim,

$$-A_2\exp[r(t^* - t)] + A_2\exp[R(t^* - t)] = 0 \Longrightarrow r = R \tag{2.55}$$

Assim, como era de se esperar, caso seja possível formar uma carteira com ações e opções que seja livre de risco, então sua rentabilidade deverá ser igual a taxa de juros de um investimento em renda fixa livre de risco.

Observe que impusemos $D^T\Theta = 0$ para obedecermos o princípio de arbitragem. Se $D^T\Theta > 0$, então pela definição 2.4, haveria oportunidade de arbitragem. Se $A^T\Theta = 0$ e $D^T\Theta < 0$ então também haveria oportunidade de arbitragem, pois se fosse escolhida uma estratégia $\Theta^* = -\Theta$, teríamos $A^T\Theta^* = 0$ e $D^T\Theta^* > 0$.

No Capítulo 5, vamos determinar as quantidades apropriadas de ação e opção para criar a carteira livre de risco e mostrar que esta carteira é apenas momentaneamente livre de risco, sendo necessários contínuos ajustes. Estes ajustes caracterizam o conceito de *hedge dinâmico*, ou seja, realização de rebalanceamento da carteira a todo instante para a proteção de seu valor.

2.3 Neutralidade a risco

Nesta seção, apresentaremos um conceito extremamente útil para a precificação de derivativos por permitir a transformação do espaço de probabilidade de forma que algumas propriedades interessantes possam ser utilizadas. Sob condições específicas, o cálculo do preço dos ativos poderá ser feito obtendo-se o valor presente do valor esperado do ativo em um momento futuro, usando como taxa de desconto ou retorno exigido, a taxa de juros livre de risco. Esta forma de precificação caracteriza um mundo denominado risco-neutro, onde todos os ativos possuem retorno esperado ou exigido igual a taxa de juros livre de risco, independente do grau de risco associado.

O mundo risco-neutro constitui-se em um modelo teórico que apresenta um resultado extremamente interessante uma vez que evita utilização do retorno esperado do ativo como taxa de desconto, como será visto a seguir. Consideremos o vetor de ativos A e a matriz de resultados D para N ativos e k estados da natureza.

$$A(t) = \begin{pmatrix} A_1(t) \\ \vdots \\ A_N(t) \end{pmatrix} \tag{2.56}$$

e

$$D = \begin{pmatrix} A_{11}(t^*) & \cdots & A_{1k}(t^*) \\ \vdots & \vdots & \vdots \\ A_{N1}(t^*) & \cdots & A_{Nk}(t^*) \end{pmatrix} \tag{2.57}$$

Novamente, vamos estabelecer que A_1 corresponde ao preço de um ativo de renda fixa livre de risco. Se supusermos que existe um equilíbrio de preços de tal forma que oportunidades de arbitragem não sejam possíveis, então pelo teorema 2.1, temos que existe $\psi > 0$ tal que:

$$A = D\psi \tag{2.58}$$

Assim,

$$A(t) = \begin{pmatrix} A_1(t) \\ \vdots \\ A_N(t) \end{pmatrix} = \begin{pmatrix} A_{11}(t^*) & \cdots & A_{1k}(t^*) \\ \vdots & \vdots & \vdots \\ A_{N1}(t^*) & \cdots & A_{Nk}(t^*) \end{pmatrix} \begin{pmatrix} \psi_1 \\ \vdots \\ \psi_k \end{pmatrix} \tag{2.59}$$

Como A_1 refere-se ao preço de um ativo livre de risco que tem taxa de juros igual a r, então:

$$A_{11}(t^*) = \cdots = A_{1k}(t^*) = A_1 \exp[r(t^* - t)] \tag{2.60}$$

Substituindo este resultado em (2.59):

$$\begin{pmatrix} A_1(t) \\ A_2(t) \\ \vdots \\ A_N(t) \end{pmatrix} = \begin{pmatrix} A_1 \exp[r(t^* - t)] & \cdots & A_1 \exp[r(t^* - t)] \\ A_{21}(t^*) & \cdots & A_{2k}(t^*) \\ \vdots & \vdots & \vdots \\ A_{N1}(t^*) & \cdots & A_{Nk}(t^*) \end{pmatrix} \begin{pmatrix} \psi_1 \\ \vdots \\ \psi_k \end{pmatrix} \tag{2.61}$$

Simplificando:

$$\begin{pmatrix} 1 \\ A_2 \\ \vdots \\ A_N(t) \end{pmatrix} = \begin{pmatrix} \exp[r(t^*-t)] & \cdots & \exp[r(t^*-t)] \\ A_{21}(t^*) & \cdots & A_{2k}(t^*) \\ \vdots & \vdots & \vdots \\ A_{N1}(t^*) & \cdots & A_{Nk}(t^*) \end{pmatrix} \begin{pmatrix} \psi_1 \\ \vdots \\ \psi_k \end{pmatrix} \tag{2.62}$$

A equação matricial acima representa o seguinte sistema de equações :

$$\begin{cases} 1 = \exp[r(t^*-t)]\psi_1 + \cdots + \exp[r(t^*-t)]\psi_k \\ A_2(t) = A_{21}\psi_1 + \cdots + A_{2k}\psi_k \\ \vdots \\ A_N(t) = A_{N1}\psi_1 + \cdots + A_{Nk}\psi_k \end{cases} \tag{2.63}$$

Sendo $\psi_i > 0$, $1 \leq i \leq k$, então podemos definir:

$$\begin{cases} P_1^A = \exp[r(t^*-t)]\psi_1 \\ \vdots \\ P_k^A = \exp[r(t^*-t)]\psi_k \end{cases} \tag{2.64}$$

tal que:

$$P_i^A > 0 \tag{2.65}$$

Além disso, pela primeira equação do sistema (2.63),

$$\sum_{i=1}^{k} P_i^A = 1 \tag{2.66}$$

Vamos estabelecer agora $\Omega_A = (e_1, \cdots, e_k)$ e definir P^A de forma que:

$$\begin{cases} P^A(e_1) = \exp[r(t^*-t)]\psi_1 \\ \vdots \\ P^A(e_k) = \exp[r(t^*-t)]\psi_k \end{cases} \tag{2.67}$$

Assim, temos um espaço de probabilidade ajustado a um mundo risco-neutro (Ω, P^A) onde o preço dos ativos pode ser definido como valor presente do valor esperado dos possíveis resultados futuros. Esta afirmação ficará clara a seguir.

Vamos multiplicar e dividir os segundos membros das equações do sistema (2.63), excluindo-se a primeira:

$$\begin{cases} 1 = \exp[r(t^*-t)]\psi_1 + \cdots + \exp[r(t^*-t)]\psi_k \\ A_2(t) = A_{21}\psi_1 \frac{\exp[r(t^*-t)]}{\exp[r(t^*-t)]} + \cdots + A_{2k}\psi_k \frac{\exp[r(t^*-t)]}{\exp[r(t^*-t)]} \\ \vdots \\ A_N(t) = A_{N1}\psi_1 \frac{\exp[r(t^*-t)]}{\exp[r(t^*-t)]} + \cdots + A_{Nk}\psi_k \frac{\exp[r(t^*-t)]}{\exp[r(t^*-t)]} \end{cases} \tag{2.68}$$

Substituindo P_i^A no sistema acima:

$$\begin{cases} 1 = P_1^A + \cdots + P_k^A \\ A_2(t) = \frac{A_{21}P_1^A + \cdots + A_{2k}P_k^A}{\exp[r(t^*-t)]} \\ \vdots \\ A_N(t) = \frac{A_{N1}P_1^A + \cdots + A_{Nk}P_k^A}{\exp[r(t^*-t)]} \end{cases} \tag{2.69}$$

Podemos definir então:

$$E_{P^A(t^*)}(A_i) = \sum_{j=1}^{k} A_{ij}P_j^A \tag{2.70}$$

como a esperança, no espaço (Ω, P_A), do preço do ativo i no momento t^*, onde A_{ij} são os possíveis resultados do ativo i nos k estados da natureza e P_j^A as probabilidades ajustadas de ocorrência de cada estado da natureza.

Assim,

$$\begin{cases} 1 = P_1^A + \cdots + P_k^A \\ A_2(t) = \frac{E_{pA}(A_2)}{\exp[r(t^*-t)]} \\ \vdots \\ A_N(t) = \frac{E_{pA}(A_N)}{\exp[r(t^*-t)]} \end{cases} \tag{2.71}$$

E portanto, os preços dos ativos equivalem à esperança medida segundo um espaço de probabilidade ajustado, trazida a valor presente pela taxa de juros livre de risco. Apesar do processo de difusão depender de um parâmetro μ, conforme descrito na equação (2.6)

$$\frac{dS}{S} = \mu dt + \sigma dW \tag{2.72}$$

apenas o parâmetro de volatilidade σ_i será importante.

É importante observar que na avaliação através do mundo risco-neutro, torna-se irrelevante o retorno esperado μ de cada ativo, tendo em vista que a taxa de desconto utilizada será igual para todos: a taxa de juros livre de risco r. Caso dispuséssemos das verdadeiras probabilidades, P_R, de ocorrência de cada estado da natureza, então, poderíamos calcular os ativos da seguinte forma:

$$\begin{cases} A_1(t) = \frac{E_{pR}(A_1)}{\exp[\mu_1(t^*-t)]} \\ \vdots \\ A_N(t) = \frac{E_{pR}(A_N)}{\exp[\mu_N(t^*-t)]} \end{cases} \tag{2.73}$$

onde cada μ_i, $1 \leq i \leq N$, corresponde a uma taxa de retorno exigida pelos investidores em função do nível de risco do ativo i.

Em geral, porém, μ_i não é diretamente observável e ainda, depende do perfil de risco do investidor, isto é, sua propensão, aversão ou neutralidade ao risco. A vantagem da utilização do conceito de mundo risco-neutro, onde $\mu_1 = \cdots = \mu_N = r$, é a possibilidade da aplicação de algumas propriedades de martingais para precificar derivativos.

Assim, o mundo risco neutro corresponde a um modelo abstrato no qual todos os ativos financeiros têm a seguinte propriedade:

$$\exp(-rT)E[S_t|S_0] = S_0 \tag{2.74}$$

ou seja, o resultado médio de qualquer ativo durante o intervalo de tempo t é o valor do ativo no instante inicial ajustado pela taxa de juros livre de risco r.

Neste trabalho, porém, não aprofundaremos os estudos de martingais, considerando esta seção apenas uma análise preliminar para podermos demonstrar a fórmula de Black e Scholes através de uma outra estrutura, que será realizada no Capítulo 6.

Chapter 3

Fundamentos de cálculo estocástico

Neste capítulo, introduziremos os conceitos sobre cálculo estocástico que são necessários para a obtenção da equação diferencial de Black e Scholes. Na primeira seção serão estudadas as integrais estocásticas, inicialmente considerando-se funções-simples, chamadas de funções-escada, e posteriormente, generalizando as definições e demonstrações para funções pertencentes a M_2. Na segunda seção, apresentaremos a notação das integrais estocásticas na forma de diferenciais estocásticas e demonstraremos o lema de Itô que representa uma poderosa ferramenta de cálculo utilizada para a precificação de vários instrumentos financeiros. Finalmente, na terceira seção, serão apresentados exemplos de cálculo de integral estocástica, aplicando-se a definição e posteriormente aplicando-se o lema de Itô.

3.1 Integral estocástica

Nesta seção[1], definiremos a integral estocástica que será denotada por:

$$\int_a^b f dW_t = \int_a^b f(t) dW_t = \int_a^b f(t,\omega) dW_t(\omega) \tag{3.1}$$

com $a < b < \infty$ e $f \in M_2[a,b]$.

Primeiramente, apresentaremos a definição de integral estocástica para funções-escada pertencentes a $M_2[a,b]$ e posteriormente, generalizaremos esta definição para quaisquer funções pertencentes a $M_2[a,b]$.

Definição 3.1 *Seja f uma função-escada definida de acordo com o Apêndice. Definiremos a integral estocástica de f em relação a W_t sobre o intervalo $[a,b]$, como a variável aleatória representada por:*

$$\int_a^b f dW_t = \int_a^b f(t) dW_t = \sum_{k=0}^{n-1} f(t_k)(W_{t_{k+1}} - W_{t_k}) \tag{3.2}$$

onde $a = t_0$ e $b = t_n$.

A seguir, vamos demonstrar alguns teoremas referentes à integral estocástica definida para funções-escada.

[1] veja Arnold

Teorema 3.1 *Seja $[a,b]$ o intervalo de integração, com $a < b$. Sejam α_i números reais e f_i funções-escada tais que $f_i \in M_2[a,b]$ e $i \leq N \in \mathbb{N}$. Temos então:*

$$\int_a^b \Big[\sum_{i=1}^{N} \alpha_i f_i(t)\Big] dW_t = \sum_{i=1}^{N} \Big[\alpha_i \int_a^b f_i(t) dW_t\Big] \tag{3.3}$$

Prova:

A demonstração segue imediatamente a partir da definição dada pela equação (3.2) e pelo fato de uma combinação linear de funções-escada ser também uma função-escada com os mesmos pontos de quebra das funções originais.

Teorema 3.2 *Seja $[a,b]$ o intervalo de integração, com $a < b$. Seja f uma função-escada tal que $f \in M_2[a,b]$. Se $E|f(t)| < \infty$ para todo $t \in [a,b]$, então:*

$$E\Big[\int_a^b f(t) dW_t\Big] = 0 \tag{3.4}$$

Prova:

Vamos aplicar a definição de integral estocástica para funções-escada:

$$E\Big[\int_a^b f(t) dW_t\Big] = E\Big[\sum_{k=0}^{n-1} f(t_k)(W_{t_{k+1}} - W_{t_k})\Big] = \sum_{k=0}^{n-1} E\Big[f(t_k)(W_{t_{k+1}} - W_{t_k})\Big] \tag{3.5}$$

Sendo $f(t_k)$ progressivamente mensurável, ou não-antecipatória, e portanto independente de $(W_{t_{k+1}} - W_{t_k})$, então:

$$E\Big[\int_a^b f(t) dW_t\Big] = \sum_{k=0}^{n-1} \Big[E[f(t_k)] E[(W_{t_{k+1}} - W_{t_k})]\Big] \tag{3.6}$$

Pela propriedade do movimento browniano, Apêndice, $E[W_{t_{k+1}} - W_{t_k}] = 0$, para todo k. Portanto:

$$E\Big[\int_a^b f(t) dW_t\Big] = \sum_{k=0}^{n-1} \Big[E[f(t_k)]\ 0\Big] = 0 \tag{3.7}$$

Teorema 3.3 *Sejam $[a,b]$ o intervalo de integração, com $a < b$, e $f \in M_2[a,b]$ uma função-escada. Se $E|f(t)|^2 < \infty$, com $t \in [a,b]$, então:*

$$E\Big[\Big|\int_a^b f(t) dW_t\Big|^2\Big] = \int_a^b E|f(t)|^2 dt \tag{3.8}$$

Prova:

Vamos calcular separadamente os dois membros da equação acima, mostrando que levam a resultados iguais. Aplicando a definição de integral estocástica para funções-escada no primeiro membro da equação (3.8), temos:

$$E\Big[\Big|\int_a^b f(t) dW_t\Big|^2\Big] = E\Big[\sum_{k=0}^{n-1} f(t_k)(W_{t_{k+1}} - W_{t_k})\Big]^2$$

$$= \Big[\sum_{k=0}^{n-1}(f(t_k)(W_{t_{k+1}} - W_{t_k}))^2 + 2\sum_{k=0}^{n-1}\sum_{l>k}^{n-1}\Big(\Big(f(t_k)(W_{t_{k+1}} - W_{t_k})\Big)\Big(f(t_l)(W_{t_{l+1}} - W_{t_l})\Big)\Big)\Big] \quad (3.9)$$

A primeira parcela da equação (3.9) pode ser calculada da seguinte forma:

$$E\Big[\sum_{k=0}^{n-1}\Big(f(t_k)(W_{t_{k+1}} - W_{t_k})\Big)^2\Big] = E\Big[\sum_{k=0}^{n-1}\Big(f(t_k)^2(W_{t_{k+1}} - W_{t_k})^2\Big)\Big] =$$

$$\sum_{k=0}^{n-1} E\Big[f(t_k)^2(W_{t_{k+1}} - W_{t_k})^2\Big] \quad (3.10)$$

Sendo $f(t_k) \in M_2[a,b]$ e portanto independente de $(W_{t_{k+1}} - W_{t_k})^2$:

$$E\Big[\sum_{k=0}^{n-1}\Big(f(t_k)(W_{t_{k+1}} - W_{t_k})\Big)^2\Big] = \sum_{k=0}^{n-1}\Big[E[f(t_k)]^2 E[W_{t_{k+1}} - W_{t_k}]^2\Big] \quad (3.11)$$

Como $E|f(t)^2| < \infty$, por hipótese, e $E[W_{t_{k+1}} - W_{t_k}]^2 = (t_{k+1} - t_k)$, pela propriedade do movimento browniano dada no Apêndice, então:

$$E\Big[\sum_{k=0}^{n-1}[f(t_k)(W_{t_{k+1}} - W_{t_k})]^2\Big] = \sum_{k=0}^{n-1} E[f(t_k)^2(t_{k+1} - t_k)]$$

$$= \sum_{k=0}^{n-1} E[f(t_k)^2](t_{k+1} - t_k) \quad (3.12)$$

A segunda parcela da equação (3.9) vale zero, pois sendo $(W_{t_{k+1}} - W_{t_k})$ independente de $(W_{t_{l+1}} - W_{t_l})$, $f(t_k)$ independente de $(W_{t_{k+1}} - W_{t_k})$ e $f(t_l)$ independente de $(W_{t_{l+1}} - W_{t_l})$, temos:

$$E\Big[\sum_{k=0}^{n-1}\sum_{l>k}^{n-1}\Big((f(t_k)(W_{t_{k+1}} - W_{t_k}))(f(t_l)(W_{t_{l+1}} - W_{t_l})\Big)\Big]$$

$$= \sum_{k=0}^{n-1}\sum_{l>k}^{n-1}\Big[E\Big((f(t_k)(W_{t_{k+1}} - W_{t_k}))(f(t_l)(W_{t_{l+1}} - W_{t_l}))\Big)\Big]$$

$$= \sum_{k=0}^{n-1}\sum_{l>k}^{n-1}\Big[E[f(t_k)(W_{t_{k+1}} - W_{t_k})]\Big]\Big[E[f(t_l)(W_{t_{l+1}} - W_{t_l})]\Big]$$

$$= \sum_{k=0}^{n-1}\sum_{l>k}^{n-1}\Big[E[f(t_k)]E[(W_{t_{k+1}} - W_{t_k})]\Big]\Big[E[f(t_l)]E[(W_{t_{l+1}} - W_{t_l})]\Big] \quad (3.13)$$

Utilizando a propriedade do movimento browniano, temos: $E[W_{t_{k+1}} - W_{t_k}] = E[W_{t_{l+1}} - W_{t_l}] = 0$ e portanto:

$$E\Big[\sum_{k=0}^{n-1}\sum_{l>k}^{n-1}\Big((f(t_k)(W_{t_{k+1}} - W_{t_k}))(f(t_l)(W_{t_{l+1}} - W_{t_l}))\Big)\Big] = 0 \quad (3.14)$$

Substituindo (3.12) e (3.14) em (3.9), temos que o primeiro membro da equação (3.8) é igual a:

$$E\left[\left|\int_a^b f(t)dWt\right|^2\right] = \sum_{k=0}^{n-1} E\left[f(t_k)^2\right](t_{k+1}-t_k) \tag{3.15}$$

Vamos agora calcular o segundo membro da equação (3.8). Considerando que $f(t)$ pode ser escrita como:

$$f(t) = \sum_{k=0}^{n-1} f(t_k)\chi_{[t_k,t_{k+1}]} \tag{3.16}$$

onde χ é a função indicadora, então:

$$\int_a^b E|f(t)|^2 dt = E\int_a^b |f(t)|^2 dt = E\left[\sum_{k=0}^{n-1}\int_{t_k}^{t_{k+1}}[f(t_k)]^2 dt\right] \tag{3.17}$$

Integrando em relação a t:

$$\int_a^b E|f(t)|^2 dt = E\left[\sum_{k=0}^{n-1}[f(t_k)]^2(t_{k+1}-t_k)\right] = \sum_{k=0}^{n-1} E\left[f(t_k)^2(t_{k+1}-t_k)\right]$$

$$= \sum_{k=0}^{n-1} E\left[f(t_k)^2\right](t_{k+1}-t_k) \tag{3.18}$$

Comparando, temos (3.15) = (3.18) e portanto:

$$E\left[\left|\int_a^b f(t)dW_t\right|^2\right] = \int_a^b E|f(t)|^2 dt \tag{3.19}$$

Lema 3.1 *Seja $f \in M_2[a,b]$ uma função-escada. Então, para quaisquer $N > 0$ e $c > 0$, temos:*

$$P\left[\left|\int_a^b f(t)dW_t\right| > 0\right] \leq \frac{N}{c^2} + P\left[\int_a^b |f(t)|^2 dt > N\right] \tag{3.20}$$

Prova:

Suponha que $f(t) = f(t_k)$ para $t_k \leq t < t_{k+1}$, onde $a = t_0 < t_1 < \ldots < t_n = b$. Vamos construir uma função auxiliar f_N definida da seguinte forma:

$$f_N(t) = \begin{cases} f(t) & \text{se } \int_a^{t_{k+1}} |f(t)|^2 dt \leq N,\ t_k \leq t < t_{k+1} \\ 0 & \text{se } \int_a^{t_{k+1}} |f(t)|^2 dt > N,\ t_k \leq t < t_{k+1} \end{cases} \tag{3.21}$$

A função f_n é também, por construção, uma função-escada pertencente a $M_2[a,b]$, pois $\int_a^{t_{k+1}} |f(t)|^2 dt$ é F_t-mensurável.

Também por construção, temos:

$$\int_a^b |f(t_n)|^2 dt = \sum_{k=0}^{n-1} |f_N(t_k)|^2(t_{k+1}-t_k) \leq N \tag{3.22}$$

Sendo dado que $|f_n(t_k)|^2 \geq 0$ e $(t_{k+1}-t_k) > 0$, então individualmente cada parcela do somatório acima também é menor que N, ou seja:

$$|f_N(t_k)|^2(t_{k+1}-t_k) \leq N \Longrightarrow |f_N(t_k)|^2 \leq \frac{N}{t_{k+1}-t_k} < \infty \tag{3.23}$$

Portanto,

$$E\Big[|f_n(t_k)|^2\Big] \leq \infty \tag{3.24}$$

Utilizando o teorema 3.3 e aplicando o resultado da equação anterior, temos

$$E\Big[\Big|\int_a^b f_N(t)dW_t\Big|^2\Big] = \int_a^b E|f_N(t)|^2dt \leq N \tag{3.25}$$

Relembrando a construção de f_N, temos que $f_N \neq f$ se e somente se $\int_a^b |f(t)|^2dt > N$. Assim:

$$P\Big[\sup_{a\leq t\leq b} |f_N(t) - f(t)| > 0\Big] = P\Big[\int_a^b |f(t)|^2dt > N\Big] \tag{3.26}$$

Vamos agora definir as seguintes variáveis α e β:

$$\alpha = \int_a^b f_N(t)dW_t$$

$$\beta = \int_a^b (f(t) - f_N(t))dW_t$$

Aplicando a desigualdade triangular, Apêndice, temos:

$$|\alpha + \beta| \leq |\alpha| + |\beta| \tag{3.27}$$

Então:

$$\Big|\int_a^b f_N(t)dW_t + \int_a^b (f(t) - f_N(t))dW_t\Big| \leq \Big|\int_a^b f_N(t)dW_t\Big| + \Big|\int_a^b \Big(f(t) - f_N(t)\Big)dW_t\Big| \tag{3.28}$$

Assim:

$$\Big|\int_a^b f(t)dW_t\Big| \leq \Big|\int_a^b f_N(t)dW_t\Big| + \Big|\int_a^b \Big(f(t) - f_N(t)\Big)dW_t\Big| \tag{3.29}$$

Aplicando a função probabilidade aos dois termos da inequação (3.29):

$$P\Big[\Big|\int_a^b f(t)dW_t\Big| > c\Big] \leq P\Big[\Big|\int_a^b f_N(t)dW_t\Big| + \Big|\int_a^b \Big(f(t) - f_N(t)\Big)dW_t\Big| > c\Big] \tag{3.30}$$

Sabemos, pelo Apêndice, que

$$P[A + B] \geq P[A] + P[B] \tag{3.31}$$

Aplicando a propriedade acima na inequação (3.30):

$$P\Big[\Big|\int_a^b f(t)dW_t\Big| > c\Big] \leq P\Big[\Big|\int_a^b f_N(t)dW_t\Big| > c\Big] + P\Big[\Big|\int_a^b \Big(f(t) - f_N(t)\Big)dW_t\Big| > c\Big] \tag{3.32}$$

Sendo $c > 0$, então:

$$P\Big[\Big|\int_a^b \Big(f(t) - f_N(t)\Big)dW_t\Big| > 0\Big] \geq P\Big[\Big|\int_a^b \Big(f(t) - f_N(t)\Big)dW_t\Big| > c\Big] \tag{3.33}$$

e portanto,

$$P\Big[\Big|\int_a^b f(t)dW_t\Big| > c\Big] \leq P\Big[\Big|\int_a^b f_N(t)dW_t\Big| > c\Big] + P\Big[\Big|\int_a^b \Big(f(t) - f_N(t)\Big)dW_t\Big| > 0\Big] \tag{3.34}$$

Como

$$P\Big[\Big|\int_a^b f_N(t)dW_t\Big|^2 > c^2\Big] = P\Big[\Big|\int_a^b f_N(t)dW_t\Big| > c\Big] \tag{3.35}$$

então:

$$P\Big[\Big|\int_a^b f(t)dW_t\Big| > c\Big] \leq P\Big[\Big|\int_a^b f_N(t)dW_t\Big|^2 > c^2\Big] + P\Big[\Big|\int_a^b \Big(f(t) - f_N(t)\Big)dW_t| > 0\Big] \tag{3.36}$$

Aplicando a desigualdade de Chebyshev apresentada no Apêndice, temos:

$$P\Big[\Big|\int_a^b f_N(t)dW_t\Big|^2 > c^2\Big] = \frac{E\Big[|\int_a^b f_N(t)dW_t|^2\Big]}{c^2} \tag{3.37}$$

e substituindo (3.37) em (3.36):

$$P\Big[\Big|\int_a^b f(t)dW_t\Big| > c\Big] \leq \frac{E\Big[|\int_a^b f_N(t)dW_t|^2\Big]}{c^2} + P\Big[\Big|\int_a^b (f(t) - f_N(t))dW_t\Big| > 0\Big] \tag{3.38}$$

Utilizando (3.26) em (3.38):

$$P\Big[\Big|\int_a^b f(t)dW_t\Big| > c\Big] \leq \frac{E\Big[|\int_a^b f_N(t)dW_t|^2\Big]}{c^2}\Big] + P\Big[\int_a^b |f(t)|^2 dt > N\Big] \tag{3.39}$$

Substituindo (3.25) em (3.39):

$$P\Big[\Big|\int_a^b f(t)dW_t\Big| > c\Big] \leq \frac{N}{c^2}\Big] + P\Big[\int_a^b |f(t)|^2 dt > N\Big] \tag{3.40}$$

Lema 3.2 *Sejam $f \in M_2[a,b]$ e $f_n \in M_2[a,b]$ uma sequência de funções-escada, tal que f_n converge em probabilidade para f, ou seja,*

$$\int_a^b |f(t) - f_n(t)|^2 \xrightarrow{P} 0 \tag{3.41}$$

então

$$\int_a^b f_n(t)dW_t \xrightarrow{P} I(f) \tag{3.42}$$

onde $I(f)$ é uma variável aleatória que não depende de uma escolha específica da sequência f_n.

Prova:
Consideremos f_m uma outra sequência de funções tal que:

$$\int_a^b |f(t) - f_m(t)|^2 \xrightarrow{P} 0 \tag{3.43}$$

Vamos utilizar a desigualdade apresentada no Apêndice, fazendo $A = f_n$, $B = f_m$ e $C = f$. Assim, temos:

$$\int_a^b |f_n(t) - f_m(t)|^2 dt \leq 2\int_a^b |f(t) - f_n(t)|^2 dt + 2\int_a^b |f(t) - f_m(t)|^2 dt \tag{3.44}$$

Pela hipótese (3.41) e pela definição da sequência em (3.43), quando $n, m \to \infty$:

$$\int_a^b |f_n(t) - f_m(t)|^2 dt \xrightarrow{P} 0 \tag{3.45}$$

Pela definição de convergência em probabilidade, a equação (3.45) significa que para todo $\epsilon > 0$:

$$\lim_{n,m\to\infty} P\left[\left|\int_a^b \Big(f_n(t) - f_m(t)\Big) dW_t\right|^2 > \epsilon\right] = 0 \tag{3.46}$$

Utilizando o lema anterior, considerando $f_n - f_m$:

$$\lim_{n,m\to\infty} \sup P\left[\left|\int_a^b \Big(f_n(t) - f_m(t)\Big) dW_t\right| > \delta\right] \leq \frac{\epsilon}{\delta^2} + \lim_{n,m\to\infty} \sup P\left[\int_a^b |f_n(t) - f_m(t)|^2 dt > \epsilon\right] \tag{3.47}$$

Sabendo-se (3.46) e substituindo em (3.47), temos:

$$\lim_{n,m\to\infty} \sup P\left[\left|\int_a^b \Big(f_n(t) - f_m(t)\Big) dW_t\right| > \delta\right] \leq \frac{\epsilon}{\delta^2} \tag{3.48}$$

que conduz a:

$$\lim_{n,m\to\infty} \sup P\left[\left|\int_a^b f_n(t) dW_t - \int_a^b f_m(t) dW_t\right| > \delta\right] \leq \frac{\epsilon}{\delta^2} \tag{3.49}$$

Tendo em vista que ϵ é arbitrário, podemos escolhê-lo tão pequeno de forma que:

$$\lim_{n,m\to\infty} \sup P\left[\left|\int_a^b f_n(t) dW_t - \int_a^b f_m(t) dW_t\right| > \delta\right] \leq 0 \tag{3.50}$$

Utilizando a propriedade sobre sequência de Cauchy, ou seja, uma vez que toda sequência de Cauchy estocástica também converge estocasticamente, existe uma variável aleatória $I(f)$ de forma que:

$$\int_a^b f_n(t) dW_t \xrightarrow{P} I(f) \tag{50}$$

Este limite é único, quase certamente, e é independente da escolha da sequência f_n que obedece a equação (3.41). Este fato é verdade pois se f_n e $\overline{f_n}$ são duas sequências que satisfazem (3.41), então podemos combiná-las em uma única sequência, que também obedece a equação (3.41).

Vamos agora, definir a integral estocástica para funções quaisquer, pertencentes a M_2.

Definição 3.2 *Seja $f \in M_2[a,b]$. A integral estocástica ou integral de Itô de f em W_t sobre o intervalo $[a,b]$ é definida como a variável aleatória $I(f)$ estabelecida de acordo com o lema 3.2, isto é:*

$$I(f) = \int_a^b f dW = \int_a^b f(t) dW_t \tag{3.51}$$

com

$$\int_a^b f_n(t) dW_t \xrightarrow{P} I(f) \tag{3.52}$$

e onde f_n representa uma sequência de funções-escada pertencentes a $M_2[a,b]$ que se aproxima de f da seguinte forma:

$$\int_a^b |f(t) - f_n(t)|^2 dt \xrightarrow{P} 0 \tag{3.53}$$

Para funções especiais pertencentes a $M_2[a,b]$ podemos estabelecer uma convergência para a integral estocástica mais forte do que a convergência em probabilidade, como será demonstrado no lema a seguir.

Lema 3.3 *Para toda função $f \in M_2[a,b]$ tal que:*

$$\int_a^b E|f(t)|^2 dt < \infty \tag{3.54}$$

existe uma sequência f_n de funções-escada em $M_2[a,b]$ com a mesma propriedade dada pela equação (3.54) tal que:

$$\lim_{n \to \infty} \int_a^b E|f_n(t) - f(t)|^2 dt = 0 \tag{3.55}$$

e portanto

$$\int_a^b f_n(t) dW_t \xrightarrow{L^2} \int_a^b f(t) dW_t \tag{3.56}$$

Prova:
Considerando o lema 3.2, sempre existe uma sequência $\overline{f_n}$ de funções-escada tal que

$$\int_a^b |f(t) - \overline{f_n}(t)|^2 dt \xrightarrow{P} 0 \tag{3.57}$$

Vamos construir uma função auxiliar g_N definida da seguinte forma:

$$g_N(x) = \begin{cases} x & \text{se } |x| \leq N \\ 0 & \text{se } |x| > N \end{cases} \tag{3.58}$$

A função g_N tem três propriedades imediatas. Dados x e $y \in \mathbb{R}$
a) $|g_N(x) - g_N(y)| \leq |x - y|$
b) $\sup |x - y| \leq 2N$
c) $|g_N(f(t)) - f(t)|^2 \leq |f(t)|^2$

Aplicando a função g_N para $f(t)$ e $\overline{f_n}(t)$ e utilizando a propriedade (a), temos:

$$|g_N(f(t)) - g_N(\overline{f_n}(t))| \leq |f(t) - \overline{f_n}(t)| \tag{3.59}$$

Elevando a equação (3.59) ao quadrado e integrando os dois membros em relação a t, obtemos:

$$\int_a^b |g_N(f(t)) - g_N(\overline{f_n}(t))|^2 dt \leq \int_a^b |f(t) - \overline{f_n}(t)|^2 dt \tag{3.60}$$

A partir de (3.57), temos:

$$\int_a^b |g_N(f(t)) - g_N(\overline{f_n}(t))|^2 dt \leq \int_a^b |f(t) - \overline{f_n}(t)|^2 dt \xrightarrow{P} 0 \tag{3.61}$$

Utilizando a propriedade (b):

$$\int_a^b |g_N(f(t)) - g_N(\overline{f_n}(t))|^2 dt \leq \int_a^b |2N|^2 dt = 4N^2(a-b) \tag{3.62}$$

e pelo teorema da convergência dominada, Apêndice, considerando cada evento $\omega \in \Omega$, segue-se que, quando $n \to \infty$:

$$E\left[\int_a^b |g_N(f(t)) - g_N(\overline{f_n}(t))|^2 dt\right] = \int_a^b E\left[|g_N(f(t) - g_N(\overline{f_n}(t))|^2\right] dt \to 0 \tag{3.63}$$

Aplicando novamente o teorema da convergência dominada, considerando agora a variável $(t, \omega) \in [a, b] \times \Omega$ e utilizando a propriedade (c), quando $N \to \infty$:

$$\int_a^b E\left[|g_N(f(t)) - f(t)|^2\right] dt \to 0 \tag{3.64}$$

Portanto, existem sequências N_k e n_k tais que, para $k \to \infty$ e utilizando a propriedade do Apêndice e o resultado (3.64):

$$\int_a^b E\left[|f(t) - g_{N_k}(\overline{f_{n_k}}(t))|^2\right] dt \leq$$

$$2\int_a^b E\left[|f(t) - g_{N_k}(f(t))|^2\right] dt + 2\int_a^b E\left[|g_{N_k}(f(t)) - g_{N_k}(\overline{f_{n_k}}(t))|^2\right] dt \to 0 \tag{3.65}$$

Portanto, podemos escolher:

$$f_k(t) = g_{N_k}(\overline{f_{n_k}}(t)) \tag{3.66}$$

de tal forma que obtenhamos uma sequência de funções-escada tais que:

$$\lim_{k \to \infty} \int_a^b E[|f_k(t) - f(t)|^2] dt = 0 \tag{3.67}$$

Assim, para esta sequência temos:

$$\int_a^b f_k dW_t \xrightarrow{P} \int_a^b f dW_t \tag{3.68}$$

Porém, em virtude do teorema 3.3, então, quando $k, p \to \infty$:

$$E\left[\left|\int_a^b f_k dW - \int_a^b f_p dW\right|^2\right] = \int_a^b E[|f_k - f_p|^2]dt \to 0 \tag{3.69}$$

e portanto:

$$\int_a^b f_k dW_t \xrightarrow{L^2} \int_a^b f dW_t \tag{3.70}$$

Pelo teorema do Apêndice, o conjunto de funções-escada é denso em $M_2[a,b]$. Assim, podemos estender a definição de integral estocástica para qualquer função pertencente a M_2.

As principais propriedades demonstradas anteriormente para funções-escada também são válidas para funções arbitrárias pertencentes a $M_2[a,b]$. A seguir, apresentaremos os principais resultados, cujas provas são análogas às realizadas nos itens demonstrados anteriormente.

Teorema 3.4 *Seja $[a,b]$ o intervalo de integração, com $a < b$. Sejam α_i números reais e f_i funções arbitrárias tais que $f_i \in M_2[a,b]$ e $i \leq N \in \mathbb{N}$. Temos então:*

$$\int_a^b \left[\sum_{i=1}^{N} \alpha_i f_i(t)\right] dW_t = \sum_{i=1}^{N} \left[\alpha_i \int_a^b f_i(t) dW_t\right] \tag{3.71}$$

Teorema 3.5 *Seja $[a,b]$ o intervalo de integração, com $a < b$. Seja f uma função arbitrária tal que $f \in M_2[a,b]$. Se $E|f(t)| < \infty$ para todo $t \in [a,b]$, então:*

$$E\left[\int_a^b f(t) dW_t\right] = 0 \tag{3.72}$$

Teorema 3.6 *Sejam $[a,b]$ o intervalo de integração, com $a < b$, e f uma função arbitrária tal que $f \in M_2[a,b]$. Se $E|f(t)^2| < \infty$, com $t \in [a,b]$, então:*

$$E\left[\left|\int_a^b f(t) dW_t\right|^2\right] = \int_a^b E|f(t)|^2 dt \tag{3.73}$$

Teorema 3.7 *Suponha que $f \in M_2[a,b]$ é contínua com probabilidade 1. Então:*

$$\int_a^f dW \xrightarrow{P} \sum_{k=0}^{n-1} f(t_k)(W_{t_{k+1}} - W_{t_k}) \tag{3.74}$$

onde $a = t_0 < t_1 < \cdots < t_n = b$ e $\delta_n = \max(t_{k+1} - t_k)$, $k = 0, \ldots, n-1$

Prova:

Para funções-escada não-antecipatórias, temos:

$$\int_a^b f_n dW = \sum_{k=}^{n-1} f(t_k)(W_{t_{k+1}} - W_{t_k}) \tag{3.75}$$

com

$$f_n(t) = \sum_{k=0}^{n-1} f(t_k)\chi_{[t_k,t_{k-1}]}(t) \tag{3.76}$$

De acordo com o lema 3.2, devemos mostrar que:

$$\int_a^b |f(t) - f_n(t)|^2 dt \xrightarrow{P} 0 \tag{3.77}$$

Porém a equação acima é válida, pois $f(\cdot, \omega)$ é contínua com probabilidade 1.

Teorema 3.8 *Dado $f \in M_2[a,b]$, se*

$$\int_a^b |f|^2 dt \xrightarrow{qc} 0 \tag{3.78}$$

então

$$\int_a^b f dW \xrightarrow{qc} 0 \tag{3.79}$$

Teorema 3.9 *Seja $f \in M_2[a,b]$. Se*

$$\int_a^b E|f(t)|^2 dt < \infty \tag{3.80}$$

então, para todo número c positivo:

$$P\left[\left|\int_a^b f dW\right| > c\right] \leq \frac{\int_a^b E[|f(t)|^2]dt}{c^2} \tag{3.81}$$

Observação:
Mesmo quando as funções-escada f_n são tais que:

$$\int_a^b |f_n(t) - f(t)|^2 dt \xrightarrow{qc} 0 \tag{3.82}$$

não necessariamente

$$\int_a^b f_n dW \xrightarrow{qc} \int_a^b f dW \tag{3.83}$$

Teorema 3.10 *Seja $f \in M_2[a,b]$. Suponha que:*

$$X_t = \int_a^t f(s)dW(s) \tag{3.84}$$

com $a \leq t < b$. As seguintes propriedades são válidas:
a) X_t é F_t-mensurável
b) se

$$\int_a^t E|f(s)|^2 ds < \infty \tag{3.85}$$

para todo $t \leq b$, então o par (X_t, F_t), para $t \in [a, b]$, é um martingal, ou seja: para $a \leq s \leq t \leq b$:

$$E(X_t|F_s) = X_s \tag{3.86}$$

Além disso, para $t, s \in [a, b]$, temos:

$$E[X_t] = 0 \tag{3.87}$$

e

$$E|X_t|^2 = \int_a^t E|f(u)|^2 du \tag{3.88}$$

e, para todo $c > 0$ e $a \leq \alpha \leq \beta \leq b$

$$P\Big[\sup_{\alpha \leq t \leq \beta} |X_t - X_\alpha|^2\Big] \leq 4 \int_\alpha^\beta E|f(s)|^2 ds \tag{3.89}$$

c) X_t tem funções amostrais contínuas com probabilidade 1
d) se, para um número natural k,

$$\int_\alpha^t E|f(s)|^{2k} ds < \infty \tag{3.90}$$

com $a \leq \alpha \leq t \leq b$ então:

$$E|X_t - X_\alpha|^{2k} \leq (k(2k-1))^{k-1} \int_\alpha^t E|f(s)|^{2k} ds \tag{3.91}$$

3.2 Diferencial estocástica

Até o momento, nós analisamos processos estocásticos formados por:

$$X_t(\omega) = \int_a^b f(t, \omega) dW_t(\omega) \tag{3.92}$$

Vamos agora, estudar processos estocásticos mais genéricos, com a seguinte configuração:

$$X_t(\omega) = X_a(\omega) + \int_a^b g(t, \omega) dt + \int_a^b f(t, \omega) dW_t(\omega) \tag{3.93}$$

onde
X_a é F_a-mensurável
$g(t, \cdot)$ é F_t-mensurável para todo $t \in [a, b]$ tal que:

$$\int_a^b |g(t, \omega)| dt < \infty \tag{3.94}$$

Note que o termo em $g(t)$ é pode ser calculado através da integral de Riemann. A equação acima simplesmente estabelece a definição de um processo estocástico que possui um componente aleatório representado por dW_t e um componente determinístico representado por dt.

Definição 3.3 *Estabeleceremos que o processo estocástico definido por:*

$$X_t(\omega) = X_a(\omega) + \int_a^b g(t,\omega)dt + \int_a^b f(t,\omega)dW_t(\omega) \tag{3.95}$$

possui diferencial estocástica denotada por:

$$dX_t = g(t)dt + f(t)dW_t = gdt + fdW \tag{3.96}$$

Observe que a equação (3.96) é apenas uma forma de representação da equação (3.95), uma vez que, teoricamente, W_t não é derivável.

Lema 3.4 *Lema de Itô.*
Seja $u = u(t,x)$ uma função contínua definida no intervalo $[a,b] \times I\!R$, com derivadas parciais $u_t = \frac{\partial u}{\partial t}$, $u_x = \frac{\partial u}{\partial x}$ e $u_{xx} = \frac{\partial^2 u}{\partial x^2}$ contínuas. Se X_t é um processo definido em $[a,b]$ com diferencial estocástica representada por:

$$dX_t = gdt + fdW \tag{3.97}$$

então $Y_t = u(t,X)$ possui em $[a,b]$ a diferencial estocástica:

$$dY_t = \left(u_t(t,X_t) + u_x(t,X_t)g(t) + \frac{1}{2}u_{xx}(t,X_t)f(t)^2\right)dt + u_x(t,X_t)f(t)dW_t \tag{3.98}$$

Prova:
Uma vez que o domínio de uma função-escada pode ser decomposta em um número finito de intervalos nos quais assume valores constantes (em função de t), podemos restringir nossa análise para o caso de $g(t,\omega) = g(\omega)$ e $f(t,\omega) = f(\omega)$.
Assim, o processo X_t possui a forma:

$$X_t = X_{t_0} + g(t - t_0) + f(W_t - W_{t_0}) \tag{3.99}$$

com $a = t_0 \leq t < b = T$ e onde X_{t_0}, g e f são variáveis aleatórias F_{t_0}- mensuráveis e portanto independentes de $(W_t - W_{t_0})$.
Sabendo que o processo Y_t tem a forma:

$$Y_t = u(t, X_t) \tag{3.100}$$

então

$$Y_t = u\Big(t, X_{t_0} + g(t - t_0) + f(W_t - W_{t_0})\Big) \tag{3.101}$$

com

$$Y_{t_0} = u(t_0, X_{t_0}) \tag{3.102}$$

Suponhamos agora que $a = t_0 < t_1 < ... < t_n = b = T$. Então:

$$Y_t - Y_{t_0} = \sum_{k=0}^{n-1}(u(t_{k+1}, X_{t_{k+1}}) - u(t_k, X_{t_k})) \tag{3.103}$$

Uma vez que $u(t,x)$ possui derivadas parciais conforme estabelecido na hipótese, podemos utilizar a fórmula de Taylor. Assim:

$$u(t_{k+1}, X_{t_{k+1}}) - u(t_k, X_{t_k})$$

$$= u_t(t_k + \theta_1(t_{k+1} - t_k), X_{t_k})(t_{k+1} - t_k) + u_x(t_k, X_{t_k})(X_{t_{k+1}} - X_{t_k})$$

$$+\frac{1}{2}u_{xx}(t_k, X_{t_k} + \theta_2(X_{t_{k+1}}))(X_{t_{k+1}} - X_{t_k})^2 \tag{3.104}$$

com $0 < \theta_1 < 1$ e $0 < \theta_2 < 1$. Devido à continuidade de X_t, u_t e u_{xx}, existem variáveis aleatórias α_n e β_n tais que, dado:

$$\delta_n = \max_{0 \leq k \leq n_1}(t_{k+1} - t_k) \to 0 \tag{3.105}$$

então α_n e β_n convergem para 0 com probabilidade 1 e satisfazem:

$$\max_{0 \leq k \leq n-1} |u_t(t_k + \theta_1(t_{k+1} - t_k), X_{t_k}) - u_t(t_k, X_{t_k})| \leq \alpha_n \tag{3.106}$$

$$\max_{0 \leq k \leq n_1} |u_{xx}(t_k, X_{t_k} + \theta_2(X_{t_{k+1}} - X_{t_k})) - u_{xx}(t_k, X_{t_k})| \leq \beta_n \tag{3.107}$$

Podemos portanto substituir (3.104) em (3.103). Além disso, se impusermos $\theta_1 = \theta_2 = 0$, não alteraremos o limite de $Y_t - Y_{t_0}$ em (3.103) quando $\delta_n \to \infty$ pois:

$$\sum_{k=0}^{n-1}(t_{k+1} - t_k) = t - t_0 \tag{3.108}$$

e

$$\sum_{k=0}^{n-1}(X_{t_{k+1}} - X_{t_k})^2 \xrightarrow{P} f^2(t - t_0) \tag{3.109}$$

Portanto, agora devemos mostrar que quando $\delta \to \infty$, temos:

$$\sum_{k=0}^{n-1}\left[u_t(t_k, X_{t_k})(t_{k+1} - t_k) + u_x(t_k, X_{t_k})(X_{t_{k+1}} - X_{t_k}) + \frac{1}{2}u_{xx}(t_k, X_{t_k})(X_{t_{k+1}} - X_{t_k})^2\right] \xrightarrow{P}$$

$$\int_{t_0}^{t}\left(u_t(s, X_s) + u_x(s, X_s)g + \frac{1}{2}u_{xx}(s, X_s)f^2\right)ds + \int_{t_0}^{t} u_s(s, X_s) f dW_s \tag{3.110}$$

Pelas hipóteses de continuidade temos, para $\delta_n \to \infty$, as primeiras duas somatórias da equação (3.110) podem ser obtidas:

$$\sum_{k=0}^{n-1} u_t(t_k, X_{t_k})(t_{k+1} - t_k) \xrightarrow{qc} \int_{t_0}^{t} u_t(s, X_s)ds \tag{3.111}$$

e

$$\sum_{k=0}^{n-1} u_x(t_k, X_{t_k})(X_{t_{k+1}} - X_{t_k}) \xrightarrow{P} \int_0^t u_x(s, X_s)gds + \int_{t_0}^{t} u_x(s, X_s)fdW_s \tag{3.112}$$

Vamos agora calcular a terceira somatória da equação (3.110). Utilizando (3.99) para calcular $(X_{t_{k+1}} - X_{t_k})^2$ temos:

$$\sum_{k=0}^{n-1} u_{xx}(t_k, X_{t_k})(X_{t_{k+1}} - X_{t_k})^2 =$$

$$g^2 \sum_{k=0}^{n-1} u_{xx}(t_k, X_{t_k})(t_{k+1} - t_k)^2 + 2fg \sum_{k=0}^{n-1} u_{xx}(t_k, X_{t_k})(t_{k+1} - t_k)(W_{t_{k+1}} - W_{t_k}) +$$

$$f^2 \sum_{k=0}^{n-1} u_{xx}(t_k, X_{t_k})(W_{t_{k+1}} - W_{t_k})^2 \tag{3.113}$$

As duas primeiras parcelas, devido à continuidade de u_{xx} e W_t, convergem para 0, com probabilidade 1, ou seja, para $\delta_n \to \infty$:

$$g^2 \sum_{k=0}^{n-1} u_{xx}(t_k, X_{t_k})(t_{k+1} - t_k)^2 \xrightarrow{qc} 0 \tag{3.114}$$

e

$$2fg \sum_{k=0}^{n-1} u_{xx}(t_k, X_{t_k}) \xrightarrow{qc} 0 \tag{3.115}$$

Portanto, para provarmos (3.110), devemos mostrar que, para $\delta_n \to \infty$:

$$f^2 \sum_{k=0}^{n-1} u_{xx}(t_k, X_{t_k})(W_{t_{k+1}} - W_{t_k})^2 \xrightarrow{P} \int_{t_0}^{t} \frac{1}{2} u_{xx}(s, X_s) f^2 ds \tag{3.116}$$

ou, simplificando:

$$\sum_{k=0}^{n-1} u_{xx}(t_k, X_{t_k})(W_{t_{k+1}} - W_{t_k})^2 \xrightarrow{P} \int_{t_0}^{t} \frac{1}{2} u_{xx}(s, X_s) ds \tag{3.117}$$

Como, aplicando a definição de integral de Riemann, temos:

$$\lim_{\delta_n \to \infty} \sum_{k=0}^{n-1} u_{xx}(t_k, X_{t_k})(t_{k+1} - t_k) = \int_{t_0}^{t} u_{xx}(s, X_s) ds \tag{3.118}$$

então, devemos mostrar que

$$S_n \xrightarrow{P} 0 \tag{3.119}$$

e portanto, M é uma variável aleatória finita quase certamente, ou seja, o segundo membro da equação (3.128) pode ser transformado em um número arbitrariamente pequeno, através da escolha de um N suficientemente grande.

Como

$$P[|S_n| > \epsilon] \leq P[|S_n^N| > \epsilon] + P[S_n \neq S_n^N] \tag{3.130}$$

então

$$S_n \xrightarrow{P} 0 \tag{3.131}$$

e, portanto, provamos que:

$$dY_t = \left(u_t(t, X_t) + u_x(t, X_t)g(t) + \frac{1}{2}u_{xx}(t, X_t)f(t)^2\right)dt + u_x(t, X_t)f(t)dW_t \tag{3.132}$$

Observação:

O lema de Itô terá papel fundamental na obtenção da equação diferencial de Black e Scholes. Devido à sua construção, o lema de Itô é uma ferramenta para a representação de funções aleatórias através de diferenciais estocásticas.

Como já foi enfatizado, a diferencial estocástica é apenas uma notação que visa simplificar a integral estocástica. Assim, a própria definição de diferencial estocástica possibilita o cálculo da integral estocástica. No cálculo tradicional, determinístico, o teorema fundamental do cálculo interliga as definições, inicialmente independentes, de derivada e integral. A resolução de uma equação diferencial implica a resolução de uma equação integral correspondente devido à validade do teorema fundamental do cálculo. No cálculo estocástico, a diferencial estocástica é definida a partir da integral. Com isso, a resolução de equação diferencial estocástica implica naturalmente a resolução da equação integral sem a necessidade de recorrer-se a um teorema.

3.3 Aplicações práticas

Vamos apresentar exemplos onde utilizaremos os conceitos de integral e diferencial estocásticas. Aplicaremos a definição de integral estocástica e o lema de Itô para a resolução de alguns problemas. Verificaremos que o lema de Itô possibilita a simplificação dos cálculos através da aplicação de uma regra útil de diferenciação.

3.3.1 Aplicação da definição de integral estocástica

Neste exemplo, vamos calcular $\int_0^T W_t dW_t$, onde $T < \infty$, aplicando-se a definição de integral estocástica.

Vamos provar, primeiramente, que $W_t \in L^2$. Segundo o Apêndice, uma função $f(t, \omega) \in L^2$ se:

$$E\int_0^T |f(t, \omega)|^2 dt < \infty \tag{3.133}$$

Para $f(t, \omega) = W_t$:

com

$$S_n = \sum_{k=0}^{n-1} u_{xx}(t_k, X_{t_k})\Big((W_{t_{k+1}} - W_{t_k})^2 - (t_{k+1} - t_k)\Big) \tag{3.120}$$

Utilizando uma técnica de truncamento, podemos eliminar valores de u_{xx}. Assim, para $N > 0$, vamos definir:

$$I_k^N(\omega) = \begin{cases} 1 \text{ se } |X_{t_i}| \le N \text{ para todo } i \le k-1 \\ 0 \text{ caso contrário} \end{cases} \tag{3.121}$$

além de:

$$\epsilon_k = (W_{t_{k+1}} - W_{t_k})^2 - (t_{k+1} - t_k) \tag{3.122}$$

e

$$S_n^N = \sum_{k=0}^{n-1} u_{xx}(t_k, X_{t_k}) I_k^N \epsilon_k \tag{3.123}$$

Pelas propriedades do movimento browniano, Apêndice, $E[\epsilon_k] = 0$ e $E[\epsilon_k^2] = 2(t_{k+1} - t_k)^2$. Além disso, os ϵ_k são independentes entre si e são também independentes de $u_{xx}(t_k, X_{t_k}) I_k^N$. Assim, temos:

$$E[S_n^N] = 0 \tag{3.124}$$

e

$$\begin{aligned} E[(S_n^N)^2] &= \sum_{k=0}^{n-1} E[u_{xx}(t_k, X_{t_k}) I_k^N]^2 E[\epsilon_k^2] \\ &\le 2 \max_{t_0 \le t \le t, |\gamma| \le N} |u_{xx}(s, \gamma)| \sum_{k=0}^{n-1} (t_{k+1} - t_k)^2 \to 0 \end{aligned} \tag{3.125}$$

quando $\delta_n \to 0$. Portanto, para todo $N > 0$ fixo temos:

$$S_n^N \xrightarrow{L^2} 0 \tag{3.126}$$

e portanto:

$$S_n^N \xrightarrow{P} 0 \tag{3.127}$$

Considerando o erro proveniente do truncamento, podemos calculá-lo da seguinte forma:

$$P[S_n \ne S_n^N] = P\Big[\max_{t_0 \le s \le t} |X_s| > N\Big] \tag{3.128}$$

Porém:

$$\begin{aligned} M = \max_{t_0 \le s \le t} |X_s| &= \max_{t_0 \le s \le t} |X_{t_0} + g(s - t_0) + f(W_s - W_{t_0})| \\ &\le |X_{t_0}| + |g|(t - t_0) + |f| \max_{t_0 \le s \le t} |W_s - W_{t_0}| \end{aligned} \tag{3.129}$$

$$E\int_0^T |W_t|^2 dt = \int_0^T E[W_t^2]dt = \int_0^T t dt = \frac{T^2}{2} < \infty \tag{3.134}$$

Sendo W_t não antecipatória, de acordo como o Apêndice, então $W_t \in [0, T]$.

Este fato em conjunto com a equação (3.134) garante a existência da integral estocástica $\int_0^T W_t dW_t$, onde $T < \infty$.

Para calcularmos o seu valor ou, mais apropriadamente, o valor aleatório devemos seguir determinados passos:

(1) encontrar uma sequência de funções pertencentes a M_2 que convirja em L^2 para W_t;

(2) calcular a integral estocástica de cada uma das funções da sequência encontrada no passo (1)

(3) obter o limite em L^2 das integrais estocásticas calculadas no passo (2).

Este limite, obtido no passo (3) será a integral estocástica $\int_0^T W_t dW_t$.

Observe que o lema 3.2 nos garante que podemos escolher qualquer sequência que satisfaça o passo (1) já que o limite em L^2 que será calculado no passo (3) não depende de uma escolha particular desta sequência.

A sequência que utilizaremos será designada por $\{f_n(t,\omega)\}$ e terá a seguinte definição: para cada n dividimos o intervalo de $[0,T]$ em n sub-intervalos iguais e para todo $\omega \in \Omega$ definimos:

$$f_n(t,\omega) = W_{\frac{kT}{n}}(\omega) \text{ para } t \in \left[\frac{kT}{n}, \frac{(k+1)T}{n}\right) \tag{3.135}$$

Vamos verificar se $f_n(t,\omega)$ pertence a M_2. Este resultado é imediato, tendo em vista que $f_n(t,\omega)$ definido acima é progressivamente mensurável por construção. Além disso, devemos provar que $f_n(t,\omega)$ converge em L^2 para $W_t(\omega)$. Pela definição de convergência em L^2, apresentado no Apêndice, devemos mostrar que:

$$\int_0^T E\left[\left(f_n(t,\omega) - W_t(\omega)\right)^2\right]dt \to 0 \text{ quando } n \to \infty \tag{3.136}$$

Para isto, calculemos:

$$E\left[\int_0^T |f_n(t,\omega) - W_t(\omega)|^2 dt\right] = \int_0^T E\left[|f_n(t,\omega) - W_t(\omega)|^2\right]dt =$$

$$\sum_{k=0}^{n-1} \int_{\frac{kT}{n}}^{\frac{(k+1)T}{n}} E\left[\left(W_{kT/n}(\omega) - W_t(n)\right)^2\right]dt = \sum_{k=0}^{n-1} \int_{\frac{kT}{n}}^{\frac{(k+1)T}{n}} \left(\frac{kT}{n} - t\right)dt \tag{3.137}$$

Assim,

$$E\left[\int_0^T |f_n(t,\omega) - W_t(\omega)|^2 dt\right] \sum_{k=0}^{n-1} \left(\frac{kT}{n}t - \frac{t^2}{2}\right)\Big|_{\frac{kT}{n}}^{\frac{(k+1)T}{n}}$$

$$= \sum_{k=0}^{n-1} \left[\frac{kT}{n}(k+1)\frac{T}{n} - \frac{1}{2}(k+1)^2\left(\frac{T}{n}\right)^2 - \left(\frac{kT}{n}\frac{kT}{n} - \frac{1}{2}k^2\left(\frac{T}{n}\right)^2\right)\right]$$

$$= \sum_{k=0}^{n-1} \left[\left(\frac{T}{n}\right)^2 \left(k(k+1) - \frac{1}{2}(k+1)^2 - k^2 + \frac{1}{2}k^2\right)\right] = \sum_{k=0}^{n-1} \left[-\frac{1}{2}\left(\frac{T^2}{n}\right)\right]$$

$$= -\frac{1}{2}\sum_{k=0}^{n-1}\left(\frac{T}{n}\right)^2 = -\frac{1}{2}n\left(\frac{T}{n}\right)^2 = -\frac{1}{2}\frac{T^2}{n} \tag{3.138}$$

Se fizermos $n \to \infty$, temos, com $T < \infty$:

$$\lim_{n\to\infty} E\Big[\int_0^T |f_n(t,\omega) - W_t(\omega)|^2 dt\Big] = \lim_{n\to\infty} -\frac{1}{2}\frac{T^2}{n} = 0$$

Assim, a relação (3.136) é válida e $f_n(t,\omega)$, onde f_n assume a forma dada por (3.133), representa uma sequência de funções que converge em L^2 para W_t, e portanto, o passo (1) está satisfeito.

Vamos prosseguir, calculando agora as integrais estocásticas referentes ao passo (2) discutido anteriormente. Observe que para cada n, $f_n(t,\omega)$ é uma função-escada progressivamente mensurável. Para tais funções, a definição (3.2) estabelece que suas integrais estocásticas são dadas por:

$$\begin{aligned} I_n(\omega) = \int_0^T f_n(t,\omega)dW_t(\omega) &= \sum_{k=0}^{n-1} f_n\left(\frac{kT}{n},\omega\right)\Big(W_{\frac{(k+1)T}{n}}(\omega) - W_{\frac{kT}{n}}(\omega)\Big) \\ &= \sum_{k=0}^{n-1} W_{\frac{kT}{n}}(\omega)\Big(W_{\frac{(k+1)T}{n}}(\omega) - W_{\frac{kT}{n}}(\omega)\Big) \end{aligned} \tag{3.139}$$

Para calcular a integral acima, vamos utilizar um artifício matemático simples, decorrente da propriedade do quadrado da soma apresentado no Apêndice. Sabendo-se que:

$$ab = \frac{1}{2}[(a+b)^2 - (a^2 + b^2)] \tag{3.140}$$

consideremos:

$$a = W_{\frac{kT}{n}}(\omega)$$

e

$$b = W_{\frac{(k+1)T}{n}}(\omega) - W_{\frac{kT}{n}}(\omega)$$

Então, temos, suprimindo ω na notação a seguir e aplicando a propriedade (3.140):

$$\sum_{k=0}^{n-1} W_{\frac{kT}{n}}\Big(W_{\frac{(k+1)T}{n}} - W_{\frac{kT}{n}}\Big) = \frac{1}{2}\sum_{k=0}^{n-1}\Big[\Big(W_{\frac{(k+1)T}{n}}\Big)^2 - \Big(W_{\frac{kT}{n}}\Big)^2 - \Big(W_{\frac{(k+1)T}{n}} - W_{\frac{kT}{n}}\Big)^2\Big] \tag{3.141}$$

Assim, temos:

$$\begin{aligned} I_n &= \frac{1}{2}\sum_{k=0}^{n-1}\Big[\Big(W_{\frac{(k+1)T}{n}}\Big)^2 - \Big(W_{\frac{kT}{n}}\Big)^2 - \Big(W_{\frac{(k+1)T}{n}} - W_{\frac{kT}{n}}\Big)^2\Big] \\ &= \frac{1}{2}\sum_{k=0}^{n-1}\Big[\Big(W_{\frac{(k+1)T}{n}}\Big)^2 - \Big(W_{\frac{kT}{n}}\Big)^2\Big] - \frac{1}{2}\sum_{k=0}^{n-1}\Big(W_{\frac{(k+1)T}{n}} - W_{\frac{kT}{n}}\Big)^2 \\ &= \frac{1}{2}\Big[\Big(W_{\frac{1T}{n}}\Big)^2 - W_{\frac{0T}{n}}\Big)^2 + \Big(W_{\frac{2T}{n}}\Big)^2 - W_{\frac{1T}{n}}\Big)^2 + \ldots \end{aligned} \tag{3.142}$$

$$+\left(W_{\frac{nT}{n}}\right)^2 - W_{\frac{(n-1)T}{n}}\Big)^2\Big] - \frac{1}{2}\sum_{k=0}^{n-1}\left(W_{\frac{(k+1)T}{n}} - W_{\frac{kT}{n}}\right)^2$$

$$= \frac{1}{2}\left[W_T^2 - W_0^2\right] - \frac{1}{2}\sum_{k=0}^{n-1}\left(W_{\frac{(k+1)T}{n}} - W_{\frac{kT}{n}}\right)^2 \tag{3.143}$$

O resultado acima representa a integral estocástica de f_n. Assim,

$$I_n = \frac{1}{2}\left[W_T^2 - W_0^2 - \sum_{k=0}^{n-1}\left(W_{\frac{(k+1)T}{n}} - W_{\frac{kT}{n}}\right)^2\right] \tag{3.144}$$

Tendo completado o passo (2), vamos agora ao passo (3), buscando encontrar um limite em L^2 da sequência $I_n(\omega)$. Como W_T^2 e W_0^2 são independentes de n, o limite em média de ordem 2 de (3.144), quando $n \to \infty$ depende do limite em média de ordem 2 de $\sum_{k=0}^{n-1}\left(W_{\frac{(k+1)T}{n}} - W_{\frac{kT}{n}}\right)^2$.

Portanto, vamos encontrar o limite em L^2 da somatória acima. Para isto, consideremos como candidato a este limite, o valor $E\left[\sum_{k=0}^{n-1}\left(W_{\frac{(k+1)T}{n}} - W_{\frac{kT}{n}}\right)^2\right]$

A esperança acima pode ser calculada da seguinte forma:

$$E\left[\sum_{k=0}^{n-1}\left(W_{\frac{(k+1)T}{n}} - W_{\frac{kT}{n}}\right)^2\right] = \sum_{k=0}^{n-1} E\left(W_{\frac{(k+1)T}{n}} - W_{\frac{kT}{n}}\right)^2$$

$$= \sum_{k=0}^{n-1}\left[\frac{(k+1)T}{n} - \frac{kT}{n}\right] = \left[\frac{1T}{n} - \frac{0T}{n}\right] + \left[\frac{2T}{n} - \frac{1T}{n}\right] + \ldots + \left[\frac{nT}{n} - \frac{(n-1)T}{n}\right] = T - 0 = T \tag{3.145}$$

Se o valor candidato T for realmente o limite em média de ordem 2 de $\sum_{k=0}^{n-1}\left(W_{\frac{(k+1)T}{n}} - W_{\frac{kT}{n}}\right)^2$, então devemos ter, quando $n \to \infty$:

$$E\left[\sum_{k=0}^{n-1}\left(W_{\frac{(k+1)T}{n}} - W_{\frac{kT}{n}}\right)^2 - T\right]^2 \to 0 \tag{3.146}$$

Vamos portanto, mostrar que (3.146) vale. Temos:

$$E\left[\sum_{k=0}^{n-1}\left(W_{\frac{(k+1)T}{n}} - W_{\frac{kT}{n}}\right)^2 - T\right]^2 =$$

$$E\left[\sum_{k=0}^{n-1}\left(W_{\frac{(k+1)T}{n}} - W_{\frac{kT}{n}}\right)^4 + 2\sum_{k=0}^{n-1}\sum_{l>k}^{n-1}\left(W_{\frac{(k+1)T}{n}} - W_{\frac{kT}{n}}\right)^2\right.$$

$$\left.\left(W_{\frac{(l+1)T}{n}} - W_{\frac{lT}{n}}\right)^2 - 2T\sum_{k=0}^{n-1}\left(W_{\frac{(k+1)T}{n}} - W_{\frac{kT}{n}}\right)^2 + T^2\right] \tag{3.147}$$

Assim,

$$E\left[\sum_{k=0}^{n-1}\left(W_{\frac{(k+1)T}{n}} - W_{\frac{kT}{n}}\right)^2 - T\right]^2 =$$

$$\sum_{k=0}^{n-1} E\Big(W_{\frac{(k+1)T}{n}} - W_{\frac{kT}{n}}\Big)^4 + 2\sum_{k=0}^{n-1}\sum_{l>k}^{n-1} E\Big[\Big(W_{\frac{(k+1)T}{n}} - W_{\frac{kT}{n}}\Big)^2\Big(W_{\frac{(l+1)T}{n}} - W_{\frac{lT}{n}}\Big)^2\Big]$$

$$-2T\sum_{k=0}^{n-1}\Big(W_{\frac{(k+1)T}{n}} - W_{\frac{kT}{n}}\Big)^2 + E[T^2] \tag{3.148}$$

Avaliando isoladamente o valor esperado de cada parcela da equação (3.148) e aplicando as propriedades do movimento browniano do Apêndice, temos:

$$E\Big[\sum_{k=0}^{n-1}\Big(W_{\frac{(k+1)T}{n}} - W_{\frac{kT}{n}}\Big)^2 - T\Big]^2 =$$

$$\sum_{k=0}^{n-1} 3\Big[\frac{(k+1)T}{n} - \frac{kT}{n}\Big]^2 + 2\sum_{k=0}^{n-1}\sum_{l>k}^{n-1}\Big[\frac{(k+1)T}{n} - \frac{kT}{n}\Big]$$

$$\Big[\frac{(l+1)T}{n} - \frac{lT}{n}\Big] - 2T\sum_{k=0}^{n-1}\Big[\frac{(k+1)T}{n} - \frac{kT}{n}\Big] + T^2 \tag{3.149}$$

Ou seja,

$$E\Big[\sum_{k=0}^{n-1}\Big(W_{\frac{(k+1)T}{n}} - W_{\frac{kT}{n}}\Big)^2 - T\Big]^2 =$$

$$= 3\sum_{k=0}^{n-1}\Big(\frac{T}{n}\Big)^2 + 2\sum_{k=0}^{n-1}\sum_{l>k}^{n-1}\Big(\frac{T}{n}\Big)^2 - 2T\sum_{k=0}^{n-1}\frac{T}{n} + T^2 =$$

$$= 3\frac{T^2}{n^2}n + 2\frac{T^2}{n^2}\frac{n(n-1)}{2} - 2T\frac{T}{n}n + T^2 =$$

$$3\frac{T^2}{n} + \frac{T^2}{n^2}n(n-1) - T^2 = 3\frac{T^2}{n} + T^2 - \frac{T^2}{n} - T^2 = 2\frac{T^2}{n} \tag{3.150}$$

Tirando o limite da equação anterior (3.150), com $n \to \infty$ e sabendo-se que $T < \infty$, temos:

$$\lim_{n\to\infty} E\Big[\sum_{k=0}^{n-1}\Big(W_{\frac{(k+1)T}{n}} - W_{\frac{kT}{n}}\Big)^2 - T\Big]^2 = \lim_{n\to\infty}\frac{2T^2}{n} = 0 \tag{3.151}$$

o que prova que $E\Big[\sum_{k=0}^{n-1}\Big(W_{\frac{(k+1)T}{n}} - W_{\frac{kT}{n}}\Big)^2\Big]$ converge em L^2 para T.

Assim, vamos encontrar o limite em L^2 da sequência $I_n(\omega)$ utilizando o resultado anterior.

$$\lim_{n\to\infty} E[I_n]^2 = \lim_{n\to\infty}\frac{1}{2}\Big[W_t^2 - W_0^2 - \sum_{k=0}^{n-1}\Big(W_{\frac{(k+1)T}{n}} - W_{\frac{kT}{n}}\Big)^2\Big] = \frac{1}{2}(W_T^2 - W_0^2) - \frac{1}{2}T \tag{3.152}$$

Lembrando que $W_0 = 0$, temos:

$$\lim_{n\to\infty} E[I_n]^2 = \frac{1}{2}(W_T^2 - T) \tag{3.153}$$

Sendo este limite, por definição, a integral estocástica, de acordo com (3.2), então:

$$I = \int_0^T W_t dW_t = \frac{1}{2}(W_T^2 - T) \tag{3.154}$$

Observação:

Note que se tivéssemos utilizado o conceito de integral de Riemann-Stieltjes, aplicando as propriedades de integração, teríamos:

$$I = \int_0^T W_t dW_t = \frac{1}{2}W_T^2\Big|_0^T = \frac{1}{2}(W_T^2 - W_0^2) = \frac{1}{2}W_T^2 \tag{3.155}$$

o que nos daria um resultado diferente do valor obtido em (3.154). Com isso, podemos observar que não é adequado utilizar as propriedades da integral de Riemann-Stieltjes para o cálculo de integral estocástica.

3.3.2 Aplicação do lema de Itô

Novamente vamos calcular $\int_0^T W_t dW_t$ como na seção anterior, porém, desta vez, utilizando o lema de Itô.

Vamos definir, de acordo com a notação utilizada em (3.97) e (3.98):

$$X_t = W_t \rightarrow dX_t = dW_t = 0dt + 1dW_t \tag{3.156}$$

Assim, temos: $g = 0$ e $f = 1$. $X = S$

Vamos fazer $Y_t = u(t, X) = \frac{1}{2}X^2$. Utilizando o lema de Itô, temos:

$$dY_t = \left(u_t(t, X_t) + u_x(t, X_t)g(t) + \frac{1}{2}u_{xx}(t, X_t)f(t)^2\right)dt + u_x(t, X_t)f(t)dW_t \tag{3.157}$$

$dC =$

onde:

$$u_t = \frac{\partial u}{\partial t} = 0 \tag{3.158}$$

$$u_X = \frac{\partial u}{\partial X} = X \tag{3.159}$$

$$u_{XX} = \frac{\partial^2 u}{\partial X^2} = 1 \tag{3.160}$$

Substituindo (3.158), (3.159) e (3.160) em (3.157), temos:

$$dY_t = \left(0 + X \times 0 + \frac{1}{2} \times 1 \times 1^2\right)dt + X \times 1dW_t = \frac{1}{2} + XdW_t \tag{3.161}$$

Como, pela equação (3.156), $X = W_t$, então:

$$dY_t = \frac{1}{2}dt + W_t dW_t \tag{3.162}$$

A equação acima, segundo as equações (3.95) e (3.96) representam o seguinte processo estocástico:

$$Y_t(\omega) = Y_0(\omega) + \int_a^b g'(t,\omega)dt + \int_a^b f'(t,\omega)dW_t(\omega) \tag{3.163}$$

onde $g' = \frac{1}{2}$ e $f' = W_t$.

Fazendo $a = 0$ e $b = T$, temos:

$$Y_T - Y_0 = \int_0^T \frac{1}{2}dt + \int_0^T W_t dW_t \tag{3.164}$$

Isolando a integral de interesse, $\int_0^T W_t dW_t$:

$$\int_0^T W_t dW_t = Y_T - Y_0 - \int_0^T \frac{1}{2}dt \tag{3.165}$$

onde, pela definição de Y_t, $Y_T = \frac{1}{2}W_T^2$ e $Y_0 = \frac{1}{2}W_0^2$.

Lembrando que também $W_0 = 0$, então:

$$\int_0^T W_t dW_t = \frac{1}{2}W_T^2 - \frac{1}{2}t\Big|_0^T = \frac{1}{2}W_T^2 - \frac{1}{2}T \tag{3.166}$$

E portanto,

$$\int_0^T W_t dW_t = \frac{1}{2}(W_T^2 - T) \tag{3.167}$$

que corresponde exatamente ao valor obtido em (3.154).

Assim, o lema de Itô, como apresentado no exemplo acima, pode simplificar o cálculo de integrais estocásticas, suprimindo a árdua aplicação da definição.

Chapter 4

Processos de mistura e suas propriedades

Neste capítulo, vamos apresentar modelos que descrevem o comportamento do preço do ativo-objeto. Tendo em vista que os derivativos, pela sua própria definição, dependem do preço de um determinado ativo-objeto, torna-se clara a importância da hipótese sobre o processo de difusão adotado. Obviamente, os modelos são teóricos e por isso, podem não corresponder exatamente ao comportamento observado no mercado. Buscaremos, entretanto, trazer modelos mais adequados à realidade do que o modelo clássico de movimento browniano geométrico.

Inicialmente, na primeira seção, vamos apresentar a construção de um processo de mistura de brownianos estebelecida através de uma variável. Investigaremos as principais formas alternativas para a construção do processo de mistura de brownianos, evidenciando a aplicabilidade para a precificação de derivativos. Como será visto na seção, algumas construções serão superiores a outras no sentido de facilitar a aplicação de procedimentos para precificação . Na segunda seção, será construído um processo de mistura formado a partir de um outro processo. Serão avaliados o processo de mistura de brownianos com a distribuição poissoniana dos momentos em que ocorrem alterações do comportamento do mercado, propiciando uma maior adequação à realidade. Adicionalmente, será analisado também uma mistura alternativa, baseada numa aproximação do processo de Poisson. Nas seções seguintes, seguirão considerações relevantes que foram obtidas através da análise dos processos de mistura construídos.

4.1 A construção da mistura de brownianos por variável

O objetivo desta sessão é construir o processo S_t que modela o comportamento do preço do ativo-objeto que fundamentará a busca por uma solução para o prêmio de uma opção. A construção desejada será aquela que confere maior facilidade na aplicação do argumento de Black e Scholes para S_t, onde S_t possui um comportamento descrito por uma mistura de brownianos. Será visto adiante que há mais de uma maneira de construir um processo cuja distribuição coincide com a de S_t. De fato, apresentaremos três maneiras. A primeira maneira (Seção 4.1.2) é compatível com as nossas exigências; a segunda (Seção 4.1.3) e a terceira (da Sessão 4.1.4) estão incluídas somente para mostrar a diversidade das possíveis construções. Iniciando nossa análise, anteriormente a estas construções, haverá na Seção 4.1.1 uma introdução ~~visando~~ à apresentação de uma visão cômoda no movimento browniano que será utilizado em seguida. c/ o objetivo de

4.1.1 Uma visão cômoda do movimento browniano

Seja $\hat{W}_t, t \in [0,T]$, o movimento browniano padrão definido num espaço probabilístico $(\hat{\Omega}, \hat{\mathcal{A}}, \hat{\mathrm{P}})$. Para um elemento $\hat{\omega} \in \hat{\Omega}$, chamaremos de *trajetória do processo* $\hat{W}_\cdot$ *no elemento* $\hat{\omega}$, ou simplesmente *trajetória*, a função $\hat{W}_\cdot(\hat{\omega})$. Chamaremos de *valor da trajetória no tempo* t o valor $\hat{W}_t(\hat{\omega})$. Designaremos por Ω o espaço de todas as possíveis trajetórias de $\hat{W}_\cdot$. Observe que segundo esta definição, se $\omega \in \Omega$, então existe $\hat{\omega} \in \hat{\Omega}$ tal que

$$\hat{W}_\cdot(\hat{\omega}) = \omega \tag{4.1}$$

A última observação nos ajuda a entender a próxima definição. Para todo $\omega \in \Omega$ e todo $t \in [0,T]$, vamos definir

$$\omega_t := \hat{W}_t(\hat{\omega}), \text{ onde } \hat{\omega} \text{ se relaciona com } \omega \text{ por (4.1)} \tag{4.2}$$

Considerando n números reais $t_1, \ldots, t_n$ e n intervalos $I_1 = [a_1, b_1], \ldots, I_n = [a_n, b_n] \subset I\!R$, com $n \in I\!N$, vamos estabelecer

$$A_{t_1,\ldots,t_n}^{I_1,\ldots,I_n} := \{\omega \in \Omega \, : \, \omega_{t_1} \in I_1, \ldots, \omega_{t_n} \in I_n\} \tag{4.3}$$

Definamos agora a probabilidade P nos conjuntos dados por (4.3) através de:

$$\mathrm{P}\Big(A_{t_1,\ldots,t_n}^{I_1,\ldots,I_n}\Big) := \hat{\mathrm{P}}\Big(\hat{\omega} \in \hat{\Omega} \, : \, \hat{W}_{t_1}(\hat{\omega}) \in I_1, \ldots, \hat{W}_{t_n}(\hat{\omega}) \in I_n\Big) \tag{4.4}$$

Designaremos por $\mathcal{A}$ a σ-álgebra gerada por conjuntos de forma (4.3) e designaremos por P a probabilidade em $\mathcal{A}$ obtida por extenção de P definida em (4.4).[1]

Definiremos o processo $W_t, t \in [0,T]$, no espaço $(\Omega, \mathcal{A}, \mathrm{P})$ por

$$W_t(\omega) := \omega_t, \text{ para todo } \omega \in \Omega \text{ e todo } t \in [0,T] \tag{4.5}$$

De acordo com nossa definição do espaço, concluiremos que $W_\cdot$ é o movimento browniano padrão. Chamaremos este processo de *movimento browniano definido no espaço de suas trajetórias* refletindo a equação (4.5).

É imediato que $\hat{W}_\cdot$ e $W_\cdot$ são processos idênticos. Porém eles estão definidos em espaços probabilísticos diferentes. A vantagem de $(\Omega, \mathcal{A}, \mathrm{P})$ sobre a $(\hat{\Omega}, \hat{\mathcal{A}}, \hat{\mathrm{P}})$ é que sua estrutura é simples e explicitamente definida, o que facilitará futuras construções e motivará a preferência na utilização do primeiro processo em detrimento do segundo.

A comodidade na utilização do espaço $(\Omega, \mathcal{A}, \mathrm{P})$ em relação ao espaço $(\hat{\Omega}, \hat{\mathcal{A}}, \hat{\mathrm{P}})$ acima, pode ser vista na construção do espaço amostral apresentado na Sessão 4.1.2 a seguir. Antecipando-nos, observe que cada elemento do espaço apresentado na seção seguinte é um par composto de uma trajetória de um browniano e um elemento de outro espaço amostral $\bar{\Omega}$. Tal par é fácil de ser composto quando o espaço de trajetórias é Ω. Vamos supor agora que $\hat{\Omega}$ substitui Ω nesta construção. Sendo que a estrutura de $\hat{\Omega}$ não é conhecida, precisaremos considerar o caso em que $\hat{\Omega}$ possui dois ou mais elementos, por exemplo, $\hat{\alpha}, \hat{\beta}$, que correspondem à mesma trajetória do browniano, ou seja, $\hat{W}_\cdot(\hat{\alpha}) = \hat{W}_\cdot(\hat{\beta})$. Neste caso, cada elemento do espaço amostral seria composto de um subconjunto de $\hat{\Omega}$ com um elemento de $\bar{\Omega}$, o que traria, pelo menos, complicações nas notações a serem utilizadas.

[1]Neste caso, como em muitos outros a seguir, definimos a probabilidade numa σ-álgebra usando o Teorema da unicidade da extenção da probabilidade de uma álgebra para a σ-álgebra gerada por esta álgebra. Este teorema, conhecida como o Teorema de Carathéodory, estabelece que se uma probabilidade $\mathrm{P}_{\mathrm{algebra}}$ foi definida numa coleção de conjuntos que formam uma álgebra, então existe uma única maneira de definir a probabilidade $\mathrm{P}_{\sigma-\mathrm{algebra}}$ na σ-álgebra gerada por essa coleção de forma tal que os valores de $\mathrm{P}_{\mathrm{algebra}}$ e de $\mathrm{P}_{\sigma-\mathrm{algebra}}$ nos conjuntos dessa coleção coincidem.

4.1.2 A construção do processo de mistura

Seja $W_t^*, t \in [0,T]$ o movimento browniano padrão definido no espaço de suas trajetórias $(\Omega^*, \mathcal{A}^*, \mathrm{P}^*)$, conforme descrito na Seção 4.1.1.

Seja $\bar{\Omega}$ o espaço discreto que contém m elementos que chamaremos $e_1, \ldots, e_m$, e seja $\bar{\mathrm{P}}$ a probabilidade definida neste espaço de forma tal que $\bar{\mathrm{P}}(e_i) = \alpha_i$. Definiremos em $\bar{\Omega}$ variáveis aleatórias $\bar{\eta}_1, \ldots, \bar{\eta}_m$ através de

$$\bar{\eta}_i(e_j) := \begin{cases} 1 & \text{se } i = j \\ 0 & \text{se } i \neq j \end{cases} \tag{4.6}$$

Agora definiremos $\Omega := \Omega^* \times \bar{\Omega}$, $\mathrm{P} := \mathrm{P}^* \times \bar{\mathrm{P}}$, $\mathcal{A} := \mathcal{A}^* \times \sigma(e_1, \ldots, e_m)$, [2] onde $\sigma(e_1, \ldots, e_m)$ é a σ-álgebra de todos os subconjuntos de $\bar{\Omega}$. Tendo em vista que $\bar{\Omega}$ é finito, então $\sigma(e_1, ..., e_m)$ é a coleção de todos os subconjuntos de $\bar{\Omega}$. Assim, através desta definição, um elemento ω de Ω será um par composto de ω^* e de e, onde o primeiro componente é um elemento de Ω^* e o segundo componente é um elemento de $\bar{\Omega}$.

Adicionalmente, um conjunto A de $\mathcal{A}$ terá uma forma genérica dada por

$$A_1^* \times \alpha \cup \ldots \cup A_n^* \times \beta \tag{4.7}$$

onde n é um número arbitrário natural, $A_1^*, \ldots, A_n^* \in \mathcal{A}^*$, e $\alpha, \ldots, \beta \in \bar{\Omega}$.

Para $\alpha \in \bar{\Omega}$, $A^* \times \alpha$ pode ser entendido como

$$\{(\omega^*, \alpha) \,:\, \omega^* \in A^*\} \tag{4.8}$$

Neste espaço probabilístico $(\Omega, \mathcal{A}, \mathrm{P})$, vamos definir as variáveis aleatórias $\eta_1, \ldots, \eta_m$ e o processo $W_t, t \in [0,T]$, de acordo com

$$\eta_i(\omega) = \eta_i(\omega^*, e) : \;=\; \bar{\eta}(e), \text{ para todo } \omega \in \Omega \text{ e todo } i \tag{4.9}$$

$$W_\cdot(\omega) = W_\cdot(\omega^*, e) : \;=\; W_\cdot^*(\omega^*), \text{ para todo } \omega \in \Omega \tag{4.10}$$

Ou seja, η_i desconsidera a primeira coordenada de ω, sendo igual ao valor que $\bar{\eta}_i$ assume na segunda coordenada de ω e $W_\cdot$ desconsidera a segunda coordenada de ω sendo igual à realização do processo $W_\cdot^*$ em sua primeira coordenada.

Uma importante relação entre $W_\cdot$ e $\{\eta_i\}$, que segue da própria definição do espaço $(\Omega, \mathcal{A}, \mathrm{P})$, é que

$$W_\cdot \text{ é independente de } \eta_1, \ldots, \eta_m \tag{4.11}$$

Assim, na verdade, a nossa escolha pela construção do espaço $(\Omega, \mathcal{A}, \mathrm{P})$ foi motivada visando a obtenção de independência dada em (4.11).

Agora definiremos o processo $S_t, t \in [0,T]$ no espaço $(\Omega, \mathcal{A}, \mathrm{P})$ através da seguinte equação :

$$S_t := S_0\Big(\eta_1 \exp(\mu_1 t + \sigma W_t) + \ldots + \eta_m \exp(\mu_m t + \sigma W_t)\Big) \tag{4.12}$$

que significa que $S_t(\omega) := S_0\Big(\eta_1(\omega) \cdot \exp(\mu_1 t + \sigma W_t(\omega)) + \ldots + \eta_m(\omega) \cdot \exp(\mu_m t + \sigma W_t(\omega))\Big)$ para todo $\omega \in \Omega$.

Segue, diretamente da construção realizada, que $S_\cdot$ dada por (4.12) tem a distribuição definida no início deste capítulo, para processos de mistura.

[2] O espaço probabilístico $(\Omega, \mathcal{A}, \mathrm{P})$ assim definido chama-se "espaço produto". Futuramente, muitos espaços serão definidos desta maneira.

4.1.3 Uma construção alternativa do processo de mistura

A construção realizada na Seção 4.1.2 utiliza uma variável aleatória para "escolher" as trajetórias do processo $S_.$. Vamos agora apresentar nesta seção uma nova construção que evita o uso de tal variável aleatória através da aplicação de uma outra idéia. Os passos desta construção são:

1. Partir de m movimentos brownianos $W^i_., i = 1, 2 \ldots, m$, definidos nos respectivos espaços de suas trajetórias $(\Omega^i, \mathcal{A}^i, \mathrm{P}^i)$.

2. Construir o espaço Ω "juntando" todos os espaços Ω^i, ou seja, $\Omega =: \cup_{i=1}^m \Omega^i$.

3. Introduzir P em Ω através da relação $\mathrm{P} = \alpha_1 \mathrm{P}^1 + \ldots \alpha_m \mathrm{P}^m$, onde cada P^i funciona apenas em sua parte Ω^i de Ω.

4. Definir $S_.$ em (Ω, P) de forma que seja igual a ω escolhida uniformemente em Ω.

Com esta construção, $S_.$ estará no espaço Ω^i com a probabilidade α_i, e, dado que foi escolhido o espaço Ω^i, a distribuição de $S_.$ coincide com a de $W^i_.$. Portanto, $S_.$ coincide em distribuição com $S_.$ dada em (4.12). Porém, a construção (4.12) fornece mais facilidade na definição da integral estocástica. Portanto, a construção de $S_.$ apresentada na seção anterior será preferida em termos de utilização.

Vamos agora formalizar a construção descrita nos passos (1.- 4.) acima. Designaremos por $W^i_t, t \in [0, T]$, $i = 1, 2, \ldots, m$, m movimentos brownianos definidos nos espaços de suas trajetórias $(\Omega^i, \mathcal{A}^i, \mathrm{P}^i), i = 1, 2, \ldots, m$. Observaremos que não é necessária a imposição de restrições com relação à dependência ou independência destes movimentos, já que este fato não afetará a distribuição de $S_.$ que será construída a seguir.

Façamos:

$$\Omega := \{\omega \, : \, \omega \in \Omega^i \text{ para algum } i \in \{1, 2, \ldots, m\}\} \tag{4.13}$$

Ou seja, Ω é simplesmente o conjunto que une todos os elementos de todos os espaços. Seria possível escrevermos $\Omega := \cup_{i=1}^m \Omega^i$, desde que impuséssemos que os espaços Ω^i não têm interseções entre si.

Denotaremos por $\mathcal{D}$ a coleção de conjuntos da seguinte forma

$$D = D^1 \cup \ldots \cup D^m, \qquad D^i \in \mathcal{A}^i, i = 1, 2, \ldots, m \tag{4.14}$$

e estabeleceremos

$$\mathrm{P}(D) := \alpha_1 \mathrm{P}^1(D^1) + \ldots + \alpha_m \mathrm{P}^m(D^m) \tag{4.15}$$

Vamos também definir

$$\mathcal{A} := \sigma(\mathcal{D}) \tag{4.16}$$

sendo P a probabilidade em $\mathcal{A}$ obtida pela extenção de P em $\mathcal{D}$ definida em (4.15).

Finalmente, seja $S_t, t \in [0, T]$, o processo definido em Ω por

$$S_t(\omega) := \omega_t, \qquad \forall \omega \in \Omega, \quad \forall t \in [0, T] \tag{4.17}$$

O processo $S_.$ então tem a mesma distribuição dos processos obtidos nas Seções 4.1.2 apresentada anteriomente e 4.1.4, que será vista a seguir.

4.1.4 Considerações sobre uma outra construção alternativa do processo de mistura

Nesta seção, vamos tecer comentários sobre a utilização do movimento browniano m-dimensional $(W^1_\cdot, \ldots, W^m_\cdot)$ na construção de $S_\cdot$, ou seja, do processo que modela o preço do ativo-objeto.

Vamos levar em consideração os seguintes fatos:

- objetivo de construção de um processo $S_\cdot$ de forma tal que ele "escolha" e "siga" um dos m movimentos brownianos,
- necessidade de obter a diferencial estocástica do processo $S_\cdot$ acima referido,
- existência da construção de diferencial estocástica para o movimento browniano m dimensional $(W^1_\cdot, \ldots, W^m_\cdot)$ [3]

Estes fatos sugerem, de forma natural, a utilização de $(W^1_\cdot, \ldots, W^m_\cdot)$ na construção do processo $S_\cdot$.

A primeira questão a ser respondida referente à execução desta construção, consiste em como "forçar" $S_\cdot$ a escolher entre as coordenadas deste movimento m-dimensional com as probabilidades desejáveis. A solução deste problema, que foi sugerida na Seção 4.1.3, não se aplica para o caso da presente seção. O motivo desta não-aplicabilidade é que $(W^1_\cdot, \ldots, W^m_\cdot)$ já está definido num espaço comum, ou seja, neste caso simplesmente não temos m espaços que podem ser "colados". Por isso a solução para o problema será seguir o procedimento da Seção 4.1.2.

Tomando $(\Omega^*, \mathcal{A}^*, \mathrm{P}^*)$ como o espaço onde $(W^1_\cdot, \ldots, W^m_\cdot)$ está definido, e repetindo os passos daquela seção, podemos construir $S_\cdot$ através de

$$S_t := S_0\Big(\eta_1 \exp(\mu_1 t + \sigma W^1_t) + \ldots + \eta_m \exp(\mu_m t + \sigma W^m_t)\Big) \tag{4.18}$$

Aparentemente, a fórmula (4.18) coincide com o fórmula (4.12). Porém, (4.18) define um processo m-dimensional, isto é, S_t é um vetor cuja i-ésima coodenada é $\eta_i \exp(\mu_i t + \sigma W^i_t)$. Uma vez que precisamos de um processo unidimensional, teremos que projetar todas as coordenadas de S_t de (4.18) num mesmo espaço. A definição (4.18) junto com a posterior projeção dos componentes darão uma construção quase idêntica à da Seção 4.1.2. A única diferença, porém, consiste no fato de que nesta seção utilizamos m movimentos brownianos independentes enquanto na Seção 4.1.2 usamos um único movimento browniano. Tendo em vista que a utilização de m brownianos não facilita o alcance dos nossos objetivos, em especial, o cálculo de integral estocástica, novamente preferiremos a construção do processo de mistura realizada na Seção 4.1.2.

4.2 A construção da mistura de brownianos por processo

Nas seções anteriores construímos o processo de mistura que escolhe e segue durante todo intervalo de tempo $[0, T]$, um dos m movimentos brownianos possíveis. Nesta seção, definiremos o processo $S_\cdot$, que escolherá um dos m possíveis movimentos brownianos em vários pontos do intervalo $[0, T]$, e que seguirá o movimento escolhido até o próximo momento de escolha. Esses momentos ou instantes de escolha serão determinados por um Processo Pontual Poisson, e a cada ponto será vinculada uma

[3] veja Arnold

variável aleatória cujo valor determinará a escolha neste ponto. Vamos estabelecer que as variáveis aleatórias vinculadas aos diferentes pontos serão independentes e identicamente distribuídas.

Designaremos por $P^*_t, t \geq 0$, um processo pontual qualquer. Este processo, conforme características específicas, será denominado posteriormente de Processo Pontual de Poisson ou Processo Pontual de Poisson Aproximado. Vamos denotar por $(\Omega^*, \mathcal{A}^*, \mathrm{P}^*)$ o espaço no qual este processo está definido. A realização de $P^*_\cdot$ em $\omega^* \in \Omega^*$ é uma sequência crescente de valores reais, $x_1, x_2, \ldots$, denominados *pontos*, caracterizando a nomenclatura de *Processo Pontual* ao processo $P^*_\cdot$. Vamos definir uma sequência de variáveis aleatórias $\{X^*_i\}_{i=1}^{\infty}$ sendo que X^*_i corresponde ao i-ésimo ponto de $P^*_\cdot$. Tendo em vista que estas variáveis aleatórias são determinadas por $P^*_\cdot$, então correspondem a variáveis aleatórias definidas no espaço $(\Omega^*, \mathcal{A}^*, \mathrm{P}^*)$. Assim, se $\mathrm{P}^*_\cdot(\omega^*) = (x_1, x_2, \ldots, x_n, \ldots)$, então $X^*_n(\omega^*) = x_n$.

Seja $\{\bar{\eta}^i\}_{i=0}^{\infty}$ uma sequência de vetores aleatórios independentes e identicamente distribuídas de forma tal que

$$\text{para todo } i,\ \bar{\eta}^i = (\bar{\eta}^i_1, \ldots, \bar{\eta}^i_m) \text{ coincide em distribuição com } (\bar{\eta}_1, \ldots, \bar{\eta}_m) \text{ definida na sessão 4.1.2} \tag{4.19}$$

A sequência $\{\bar{\eta}^i\}$ pode ser realizada no espaço $(\bar{\Omega}^{\infty}, \bar{\mathcal{A}}^{\infty}, \bar{\mathrm{P}}^{\infty})$, onde $\bar{\Omega}^{\infty}$ significa $\bar{\Omega}_0 \times \bar{\Omega}_1 \times \ldots \times \bar{\Omega}_n \times \ldots$, e cada $\bar{\Omega}_i$ tem a mesma estrutura do espaço $\bar{\Omega}$ construído na Sessão 4.1.2.

Um elemento α de $\bar{\Omega}^{\infty}$ é um vetor com um número de coordenadas infinito $(\alpha_0, \alpha_1, \ldots)$, onde $\alpha_i \in \bar{\Omega}_i$ para todo $i \in I\!N$. O valor de $\bar{\eta}^i$, o i-ésimo vetor da sequência $\{\bar{\eta}_i\}$ em α, está determinado por α_i que é a i-ésima coordenada de α, da seguinte maneira

$$(\bar{\eta}^i_1, \ldots, \bar{\eta}^i_m)(\alpha_i) = (\bar{\eta}_1, \ldots, \bar{\eta}_m)(\alpha_i) \tag{4.20}$$

Lembremos que $\bar{\Omega}_i$ é uma cópia de $\bar{\Omega}$ e portanto, o lado direito da equação (4.20) está bem definido. Vamos agora analisar a probabilidade $\bar{\mathrm{P}}^{\infty}$. Esta probabilidade está definida em duas etapas. Na primeira etapa, $\bar{\mathrm{P}}^{\infty}$ está definida para cada conjunto na forma

$$\bar{A}_0 \times \bar{A}_1 \times \ldots \times \bar{A}_n \times \bar{\Omega}_{n+1} \times \bar{\Omega}_{n+2} \times \ldots \tag{4.21}$$

onde $\bar{A}_i \subseteq \bar{\Omega}_i, i = 0, 1, \ldots, n$. A definição de $\bar{\Omega}_i$ é então dada por

$$\bar{\mathrm{P}}^{\infty}\Big(\bar{A}_0 \times \bar{A}_1 \times \ldots \times \bar{A}_n \times \bar{\Omega}_{n+1} \times \bar{\Omega}_{n+2} \times \ldots\Big) := \bar{\mathrm{P}}[\bar{A}_0] \cdot \bar{\mathrm{P}}[\bar{A}_1] \cdot \ldots \cdot \bar{\mathrm{P}}[\bar{A}_n] \tag{4.22}$$

Novamente, lembremos que por ser $\bar{\Omega}_i$ uma cópia de $\bar{\Omega}$, $\bar{\mathrm{P}}[\bar{A}_i]$ está bem definido. Agora, na segunda etapa, a coleção de conjuntos do tipo (4.21) estende-se para uma σ-álgebra, a ser designada por $\bar{\mathcal{A}}^{\infty}$ e consequentemente, $\bar{\mathrm{P}}^{\infty}$ torna-se a probabilidade nesta σ-álgebra obtida pela extenção da probabilidade definida em (4.22).[4]

Precisamos ainda de mais um processo e de mais um espaço probabilístico: o processo $W^*_\cdot$ e o espaço de suas trajetórias $(\Omega^*, \mathcal{A}^*, \mathrm{P}^*)$ definidas como na Seção 4.1.2.

Vamos definir o espaço probabilístico

$$\Omega := \Omega^* \times \bar{\Omega}^{\infty} \times \Omega^*, \quad \mathcal{A} := \mathcal{A}^* \times \bar{\mathcal{A}}^{\infty} \times \mathcal{A}^*, \quad \mathrm{P} := \mathrm{P}^* \times \bar{\mathrm{P}}^{\infty} \times \mathrm{P}^* \tag{4.23}$$

[4]A construção do produto infinito de espaços de probabilidade, feita aqui, é parecida com a construção da Seção 4.1.2. Porém, exige-se um certo cuidado em virtude da infinidade de componentes neste produto. A afirmação de que tal construção é correta segue do Teorema de Kolmogorov. Veja, por exemplo, Shiryayev

Neste espaço, estabeleceremos também um processo de Poisson pontual $P.$, uma sequência de variáveis aleatórias $\{X_i\}_{i=1}^{\infty}$, uma sequência de vetores aleatórios $\{\eta^i\}_{i=1}^{\infty} = \{\eta_1^i, \ldots, \eta_m^i)\}_{i=1}^{\infty}$, e um movimento browniano $W.$. As definições são, para todo $\omega \in \Omega$, todo i e todo j, dadas a seguir.

$$
\begin{aligned}
P.(\omega) = P.(\omega^*; (\alpha_0, \alpha_1, \ldots, \alpha_i, \ldots), \omega^*) : &= \bar{P}.(\omega^*), & (4.24)\\
X_i(\omega) = X_i(\omega^*; (\alpha_0, \alpha_1, \ldots, \alpha_i, \ldots), \omega^*) : &= \bar{X}_i(\omega^*), & (4.25)\\
\eta_j^i(\omega) = \eta_j^i(\omega^*; (\alpha_0, \alpha_1, \ldots, \alpha_i, \ldots), \omega^*) : &= \bar{\eta}_j^i(\alpha_i), & (4.26)\\
W.(\omega) = W.(\omega^*; (\alpha_0, \alpha_1, \ldots, \alpha_i, \ldots), \omega^*) : &= W.^*(\omega^*) & (4.27)
\end{aligned}
$$

Observe, que (4.24), (4.25), (4.26) e (4.27) valem por analogia com (4.9) e (4.10). Portanto, aplica-se neste caso também a observação feita imediatamente após (4.10).

No espaço probabilístico $(\Omega, \mathcal{A}, \mathrm{P})$ definiremos o processo $S_t, t \in [0, T]$, da seguinte maneira. Para $t \in [0, X_1(\omega)]$,

$$S_t := S_0(\eta_1^0 \exp(\mu_1 t + \sigma W_t) + \ldots + \eta_m^0 \exp(\mu_m t + \sigma W_t)) \tag{4.28}$$

Tendo definido $S_t(\omega)$ para $t \in [0, X_n(\omega)]$, definiremos agora $S_t(\omega)$ para $t \in [X_n(\omega), X_{n+1}(\omega)]$:

$$S_t := S_{X_n}(\eta_1^n \exp(\mu_1 t + \sigma W_t) + \ldots + \eta_m^n \exp(\mu_m t + \sigma W_t)) \tag{4.29}$$

4.2.1 O processo de mistura com a distribuição poissoniana dos instantes de mudanças

O processo de mistura com a distribuição Poissoniana dos instantes de mudanças seria o nome do processo $S.$ definido na sessão anterior, quando $P_t, t \geq 0$, é um Processo Pontual de Poisson. Esse processo será abreviado por MP. Se precisarmos enfatizar que a intensidade do processo é λ, então escreveremos MP(λ).

4.2.2 O processo de mistura com a distribuição poissoniana aproximada dos instantes de mudanças

O primeiro passo é construir um processo pontual que seja uma boa aproximação do Processo Pontual de Poisson com a intensidade λ e será abreviado por PPP(λ). O processo que aproxima será chamado Processo Pontual de Poisson Aproximado com a intensidade λ e abreviado por PPPA(λ). O processo de mistura no qual a distribuição dos instantes de mudanças segue o PPPA(λ) será abreviado por MPA(λ), ou simplesmente MPA.

A aproximação que utilizaremos segue o estudo de Galves e Ferrari [5], sendo consequência da construção do PPP cujo procedimento é apresentado abaixo.

(1) Dividiremos o eixo do tempo em intervalos $[t_i, t_{i+1}], i \in \mathbb{N}$ e designaremos por $\mathcal{D}_i$ o intervalo $[t_i, t_{i+1}]$.

(2) Realizaremos em cada intervalo $\mathcal{D}_i$ uma variável aleatória D_i que tem distribuição de Poisson com parâmetro $\lambda|\mathcal{D}_i|$, onde $|\mathcal{D}_i|$ é o comprimento do intervalo $\mathcal{D}_i$. Assim,

[5] veja também James

$$P[D_i = k] = \exp(-\lambda|\mathcal{D}_i|)\frac{(\lambda|\mathcal{D}_i|)^k}{k!}, \qquad k = 0, 1, \ldots$$

sendo que as variáveis aleatórias $\{D_i\}_{i \in I\!N}$ são independentes pela nossa definição.

(3) Em cada intervalo $\mathcal{D}_i$ escolheremos D_i pontos independentemente, cada um conforme a distribuição uniforme neste intervalo.

É conhecido [6] que o conjunto dos pontos obtido aplicando-se os passos *(1)-(3)* coincide em distribuição com o PPP com a intensidade λ.

Agora observaremos, que se $|\mathcal{D}_i|$ é suficientemente pequeno, então

$$P[D_i = 1] \approx \lambda|\mathcal{D}_i| \tag{4.30}$$

$$P[D_i = 0] \approx 1 - \lambda|\mathcal{D}_i| \tag{4.31}$$

$$P[D_i \geq 2] \approx 0 \tag{4.32}$$

As equações acima (4.30), (4.31) e (4.32) são usadas para a obtenção do PPPA. Primeiramente, escolheremos os intervalos $\mathcal{D}_i$, $i \in I\!N$. Esta escolha é coerente com o tipo de dados amostrais disponíveis no mercado financeiro, Tecnicamente, toda negociação com a ação tem seu preço divulgado no mercado e portanto é passível de fazer parte da amostra coletada. Façamos então: t_{2i} o momento de abertura do pregão do i-ésimo dia e t_{2i+1} o momento de fechamento do pregão do i-ésimo dia. Se considerarmos que o pregão fica aberto durante 8 horas e fechado durante 16 horas em um dia e que todos os dias haja pregão, então, consequentemente, por exemplo, podemos ter $\mathcal{D}_{2i} = \frac{1}{8}\frac{1}{365}$ e $\mathcal{D}_{2i+1} = \frac{1}{15}\frac{1}{365}$, correspondendo a uma base diária de 24 horas. Observe que, com o objetivo de simplificar a exposição, não foi modelado o comportamento de dias úteis, ano bissextos nem eventuais interrupções que possam ocorrer no mercado, como por exemplo, o recém instituído *circuit breaker*. Obviamente, os ajustes em função dos intervalos de tempo considerados são imediatos.

O processo PPPA com a intensividade λ será definido então como um processo pontual no qual os pontos são colocados conforme o seguinte procedimento: em cada untervalo $[i, i+1]$ um ponto será escolhido conforme a distribuição uniforme em $[i, i+1]$ se $Q_i = 1$ e nenhum ponto será escolhido $[i, i+1]$, se $Q_i = 0$, onde $\{Q_i\}_{i \in I\!N}$ são variáveis aleatórias independentes e identicamente distribuídas de acordo com

$$P[Q_i = 1] = \lambda|\mathcal{D}_i|, \qquad P[Q_i = 0] = 1 - \lambda|\mathcal{D}_i|$$

Pela discussão acima, o PPPA construído é uma aproximação do PPP com a intensidade λ, assumindo que o valor de λ é suficientemente pequeno para que as aproximações dadas por (4.30), (4.31) e (4.32) sejam adequadas.

Agora, vamos construir o processo de mistura de brownianos com a distribuição poissoniana aproximada dos instantes de mudanças do comportamento do mercado. Este processo será denotado por MPA. Neste processo, os instantes de mudança de cenários são pontos de um PPPA. Em cada ponto será escolhido um dos m possíveis cenários, sendo que a probabilidade de escolher o i-ésimo cenário é α_i e todas as escolhas são independentes. Suponha que s é um ponto do PPPA e t o próximo ponto

[6] veja Galves e Ferrari

desse processo. Seja S_s o valor da ação no tempo s, então se o i-ésimo cenário foi escolhido no instante s, teremos que

$$S_{s+\tau} = S_s \exp\{\mu_i \tau + \sigma_i W^i_\tau\}, \qquad \tau \in [0, t-s] \tag{4.33}$$

onde escolher um cenário significa escolher um dos pares $(\mu_1, \sigma_1), \ldots, (\mu_m, \sigma_m)$.

Se S_0 é conhecido, então (4.33) define unicamente um processo S_t, $t \geq 0$ que será chamado de MPA.

4.3 A volatilidade do processo de mistura

Veremos adiante que a fórmula do preço de opção tem a mesma aparência tanto se $S_t = S_0 \exp\{\mu t + \sigma W_t\}$, quanto se $S.$ está sendo modelado pelo processo de mistura dado por (4.12). A diferença essencial entre esses casos está na interpretação do parâmetro σ. No primeiro caso, σ é a volatilidade do processo $S.$, no segundo caso, não necessariamente. Neste segundo caso, σ é a volatilidade de cada um dos processos $\mu_k t + \sigma W.$, $k = 1, 2, \ldots, m$, porém, isto não implica que σ represente a volatilidade do processo de mistura $S.$, como será mostrado nesta seção.

Vamos lembrar que no caso "clássico" quando $S_t = \exp(\mu t + \sigma W_t)$, o valor σ corresponde à nomenclatura de *volatilidade*. A volatilidade representa a raiz quadrada da variância do retorno durante um intervalo de tempo 1:

$$\begin{aligned} \mathrm{Var}\left[\log \frac{S_{t+1}}{S_t}\right] &= \mathrm{Var}[\mu + \sigma(W_{t+1} - W_t)] \\ &= \sigma^2 \mathrm{Var}[W_{t+1} - W_t] = \sigma^2 \end{aligned}$$

Por analogia,[7] chamaremos por a *volatilidade* do processo de mistura (4.12) à raiz quadrada do seu retorno durante um intervalo de tempo 1. Da equação (4.12) temos que

$$\log\left[\frac{S_{t+1}}{S_t}\right] = \begin{cases} \mu_1 + \sigma(W_{t+1} - W_t) & \text{se } \eta_1 = 1 \\ \vdots & \\ \mu_i + \sigma(W_{t+1} - W_t) & \text{se } \eta_i = 1 \\ \vdots & \\ \mu_m + \sigma(W_{t+1} - W_t) & \text{se } \eta_m = 1 \end{cases} \tag{4.34}$$

Daí, temos

$$\mathrm{E}\left[\log \frac{S_{t+1}}{S_t}\right] = \alpha_1 \mu_1 + \ldots + \alpha_m \mu_m \tag{4.35}$$

Nesta última passagem e nos cálculos a seguir usamos dois fatos. O primeiro fato corresponde a $\mathrm{E}(W_{t+1} - W_t) = 0$. O segundo fato é que se uma variável aleatória X é uma mistura de variáveis aleatórias $X_1, \ldots, X_m$ com os pesos $\alpha_1, \ldots, \alpha_m$, então

$$\int f(x) dF_X(x) = \sum_{i=1}^{m} \alpha_i \int f(x) dF_{X_i}(x) \tag{4.36}$$

[7] Observe, que a única razão de chamarmos a raiz quadrada da variância de (4.34) de "volatilidade" é a analogia com o caso clássico.

para f que satisfaz certas condições. Em particular, se $f(x) = x^2$ e se cada uma das variáveis aleatórias $X, X_1, \ldots, X_m$ tem segundo momento finito.

Agora, podemos calcular

$$\begin{aligned} \mathrm{E}\left[\log \frac{S_{t+1}}{S_t}\right]^2 &= \sum_{i=1}^{m} \alpha_i \mathrm{E}[\mu_i + \sigma(W_{t+1} - W_t)]^2 \\ &= \sum_{i=1}^{m} \alpha_i [\mu_i^2 + 2\mu_i \sigma \mathrm{E}(W_{t+1} - W_t) + \sigma^2 \mathrm{E}(W_{t+1} - W_t)^2] \\ &= \sum_{i=1}^{m} \alpha_i [\mu_i^2 + \sigma^2] = \sigma^2 + \sum_{i=1}^{m} \alpha_i \mu_i^2 \end{aligned} \tag{4.37}$$

Logo

$$\mathrm{Var}\left[\log \frac{S_{t+1}}{S_t}\right] = \mathrm{E}\left[\log \frac{S_{t+1}}{S_t}\right]^2 - \left(\mathrm{E}\left[\log \frac{S_{t+1}}{S_t}\right]\right)^2 = \sigma^2 + \sum_{i=1}^{m} \alpha_i \mu_i^2 - \left(\sum_{i=1}^{m} \alpha_i \mu_i\right)^2 \tag{4.38}$$

Com este procedimento, demonstramos o seguinte teorema

Teorema 1 *A volatilidade do processo de mistura definido em (4.12) é* $\sigma^2 + \sum_{i=1}^{m} \alpha_i \mu_i^2 - (\sum_{i=1}^{m} \alpha_i \mu_i)^2$.

4.4 Questões em aberto relacionadas com a construção da mistura

Seja $\{W_\cdot^{\mu,\sigma},\ \mu \in I\!R, \sigma \in I\!R^+\}$ uma família de movimentos brownianos no intervalo de tempo $[0,T]$. Seja (η, ζ). um processo com valores em $I\!R \times I\!R^+$ definido no mesmo intervalo de tempo $[0,T]$. A questão que surge consiste na possibilidade de construção do processo

$$W_t^{\eta_t, \zeta_t}, t \in [0,T] \tag{4.39}$$

Este processo comporta-se da seguinte forma: em cada instante t realiza-se uma "escolha" de parâmetros de tendência μ e de variância σ^2 através dos valores de η_t e ζ_t respectivamente, e o processo segue $W_\cdot^{\mu,\sigma}$. Porém, não está claro como realizar rigorosamente tal construção. O problema está na maneira da escolha dos parâmetros do movimento browniano a ser seguido, que muda continuamente. Um caminho que evita tal problema, é definir

$$\eta_t t + \zeta_t W_t, t \in [0,T] \tag{4.40}$$

onde $W_t, t \in [0,T]$ é um movimento browniano padrão. Intuitivamente, as distribuições dos processos (4.39) e (4.40) coincidem. Tanto esta coincidência quanto as construções rigorosas de (4.39) e (4.40) estão fora do escopo deste trabalho.

Existe uma série de perguntas sobre o processo (4.39, 4.40), as quais consideramos ser interessantes e importantes. Por exemplo, estabelecer as condições para os processos η. e ζ. que garantam que o retorno de $W_t^{\eta_t,\zeta_t}, t \in [0,T]$, seja um martingal. Outro exemplo, é obter uma maneira de inferir os parâmetros dos processos η. e ζ. através de dados amostrais do mercado financeiro.

Chapter 5

Discussão do modelo de Black e Scholes

O objetivo deste capítulo é apresentar o raciocínio desenvolvido por Black e Scholes para a precificação de opções e sua utilização para outros modelos de comportamento do preço do ativo-objeto. Na primeira seção, seguiremos o raciocínio de Black e Scholes, explicitando as passagens essenciais para a obtenção da fórmula de precificação de derivativos. Seguiremos, então, a suposição original de que o preço da ação segue um movimento browniano geométrico e portanto os conceitos de cálculo estocástico serão imprescindíveis, em especial o Lema de Itô. Na segunda seção, estabeleremos um processo estocástico para o preço da ação cujo processo de difusão possui trajetórias diferenciáveis. Vamos construir, analogamente a Black e Scholes, uma carteira constituída por ações e opções equivalente a um investimento em um ativo livre de risco. Além disso, obteremos o preço da opção através da resolução de uma equação diferencial, sem a necessidade de recorrermos aos conceitos de cálculo estocástico, devido à característica de diferenciabilidade das trajetórias. Mostraremos, porém, que o modelo de trajetórias diferenciáveis, apesar de propiciar a eliminação da aleatoriedade e portanto permitir a descrição do preço da opção através de uma equação diferencial determinística, não nos conduz a um modelo de precificação adequado, tendo em vista a sua característica de previsibilidade de valores futuros do preço do ativo-objeto. Na terceira seção, vamos estruturar uma carteira livre de risco, análoga às carteiras das seções anteriores, porém agora aplicando o conceito de hedging dinâmico para uma situação em que o processo que modela o comportamento do preço da ação é uma mistura de movimentos brownianos geométricos. Investigaremos as condições em que este novo modelo implica em uma solução para o preço do derivativo. Em particular, observaremos que quando os parâmetros σ dos componentes da mistura são iguais, uma fórmula semelhante à de Black e Scholes pode ser encontrada. A seguir, na quarta seção, vamos apresentar de forma sistematizada os principais passos do raciocínio de Black e Scholes, que foram seguidos nas seções anteriores. Esta sistematização será utilizada para podermos discutir na quinta seção uma restrição para a aplicação do raciocínio de Black e Scholes, apresentando um exemplo de precificação onde o uso indiscriminado deste raciocínio pode levar a resultados que não necessariamente são os valores corretos dos preços dos derivativos.

5.1 Precificação de opções para movimento browniano

Vamos apresentar o raciocínio utilizado por Black e Scholes para a formulação teórica do preço da opção . Consideremos que o preço da ação siga um processo estocástico representado da seguinte forma:

$$\frac{dS}{S} = \mu dt + \sigma dW \tag{5.1}$$

ou:

$$dS = \mu S dt + \sigma S dW \tag{5.2}$$

Note que S não é diferenciável, uma vez que o movimento browniano tem trajetórias que não são diferenciáveis em nenhum ponto. Assim, neste caso específico, devemos utilizar a integral estocástica para avaliarmos o preço da opção.

Vamos supor também que $C(S,t)$ seja uma função, derivável pelo menos duas vezes em relação a S e derivável pelo menos uma vez em relação a t representando o preço da opção de compra.

Sem perda de generalidade, vamos assumir que o ativo-objeto seja uma ação de uma empresa e o derivativo seja uma opção de compra ou call sobre ações que não pagarão dividendos até o vencimento.

Vamos agora construir uma carteira livre de risco, formada por quantidades balanceadas de ações e opções. Sabendo-se que uma variação no preço da ação causa uma variação no preço da call, a carteira livre de risco deve ser composta por uma posição de N_A ações e uma posição contrária de N_D calls, de forma que a variação decorrente de perdas ou ganhos na posição em ações seja totalmente compensada por ganhos ou perdas na posição em calls. A suposição de posições opostas entre ações e opções de compra decorre do fato de que $\frac{\partial C}{\partial S} > 0$, ou seja, um aumento no preço das ações causa um aumento no preço da call.

Inicialmente, o valor da carteira é portanto, considerando uma posição de N_A ações e N_D opções de compra:

$$V_P = N_A S(t) + N_D C(S,t) \tag{5.3}$$

Se uma variação no valor da ação dS causa uma variação de $\frac{\partial C}{\partial S} dS$ no valor da opção e se fizermos $N_A = 1$, ou seja, uma posição comprada em 1 ação, a eliminação do risco da carteira é obtida através da incorporação de uma posição vendida em calls dada por:

$$N_D = -\frac{1}{\frac{\partial C(S,t)}{\partial S}} \tag{5.4}$$

Observe que, pela conceituação de Black e Scholes, o valor da carteira livre de risco é imune somente a variações do preço da ação, sendo sensível, porém, à passagem do tempo. Isto é, a carteira chamada de livre de risco é apenas instantaneamente livre de risco. Adicionalmente, existe uma hipótese implícita de que a ação muda de valor em um intervalo muito pequeno de tempo e neste caso, a variação do preço da call devido a passagem do tempo é desprezível, sendo relevante apenas a variação no valor do ativo-objeto.

Substituindo os valores de N_A e N_D em (5.3). temos:

$$V_P = S(t) - \frac{1}{\frac{\partial C(S,t)}{\partial S(t)}} \tag{5.5}$$

Para variações infinitesimais do preço da ação:

$$dV_P = dS(t) - \frac{1}{\frac{\partial (S,t)}{\partial S(t)}} dC(S,t) \tag{5.6}$$

Vamos utilizar portanto o lema de Itô para calcular o prêmio da call, uma vez que $C = C(S,t)$ e $S(t)$ é função do movimento browniano, conforme (5.1). Assim, temos:

$$dC = \Big(c_t(t,S_t) + c_S(t,S_t)g(t) + \frac{1}{2}c_{SS}(t,S_t)f(t)^2\Big)dt + c_S(t,S_t)f(t)dW \tag{5.7}$$

onde

$$\begin{cases} g = \mu S \\ f = \sigma S \\ c_t = \frac{\partial C}{\partial t} \\ c_S = \frac{\partial C}{\partial S} \\ c_{SS} = \frac{\partial^2 C}{\partial S^2} \end{cases} \tag{5.8}$$

Substituindo as equações acima em (5.7):

$$dC = \Big(\frac{\partial C}{\partial t} + \frac{\partial C}{\partial S}\mu S + \frac{1}{2}\frac{\partial^2 C}{\partial S^2}\sigma^2 S^2\Big)dt + \frac{\partial c}{\partial S}\sigma S dW \tag{5.9}$$

Rearranjando as parcelas da equação acima, temos:

$$dC = \frac{\partial C}{\partial t}dt + \frac{1}{2}\frac{\partial^2 C}{\partial S^2}\sigma^2 S^2 dt + \frac{\partial C}{\partial S}(\mu S dt + \sigma S dW) \tag{5.10}$$

Sabendo (5.2), isto é , $dS = \mu S dT + \sigma S dW_t$, então:

$$dC = \frac{\partial C}{\partial t}dt + \frac{1}{2}\frac{\partial^2 C}{\partial S^2}\sigma^2 S^2 dt + \frac{\partial C}{\partial S}dS \tag{5.11}$$

Substituindo (5.11) na equação da carteira livre de risco (5.6) e simplificando a notação , $dS(t) = dS$ e $\frac{\partial C(t)}{\partial S(t)} = \frac{\partial C}{\partial S}$:

$$\begin{aligned} dV_P &= dS - \frac{1}{\frac{\partial C}{\partial S}}\Big(\frac{\partial C}{\partial t}dt + \frac{1}{2}\frac{\partial^2 C}{\partial S^2}dt + \frac{\partial C}{\partial S}dS\Big) \\ &= dS - \frac{1}{\frac{\partial C}{\partial S}}\Big(\frac{\partial C}{\partial t}dt\Big) + \frac{1}{\frac{\partial C}{\partial S}}\Big(\frac{1}{2}\frac{\partial^2 C}{\partial S^2}dt\Big) + \frac{1}{\frac{\partial C}{\partial S}}\Big(\frac{\partial C}{\partial S}dS\Big) \end{aligned} \tag{5.12}$$

Simplificando e eliminando as parcelas comums:

$$dV_P = -\frac{1}{\frac{\partial C}{\partial S}}\Big(\frac{\partial C}{\partial t}dt + \frac{1}{2}\frac{\partial^2 C}{\partial S^2}dt\Big) \tag{5.13}$$

Colocando dt em evidência:

$$dV_P = -\frac{1}{\frac{\partial C}{\partial S(t)}}\Big(\frac{\partial C}{\partial t} + \frac{1}{2}\frac{\partial^2 C}{\partial S^2}\Big)dt \tag{5.14}$$

Sabendo que esta carteira é a cada instante livre de risco, então espera-se que após um intervalo de tempo Δt, tenha uma rentabilidade igual a taxa de juros livre de risco, conforme discussão apresentada no Capítulo 2.

Com isso,

$$\Delta V_P = \Big(S - \frac{1}{\frac{\partial C}{\partial S}}C\Big)r\Delta t \tag{5.15}$$

Novamente, aproximando ΔV_P, por dV_P e Δt por dt, temos:

$$dV_P = \left(S - \frac{1}{\frac{\partial C}{\partial S}}C\right)rdt \tag{5.16}$$

Igualando-se (5.16) e (5.14):

$$\left(S - \frac{1}{\frac{\partial C}{\partial S}}C\right)rdt = -\frac{1}{\frac{\partial C}{\partial S}}\left(\frac{\partial C}{\partial t} + \frac{1}{2}\frac{\partial^2 C}{\partial S^2}\right)dt \tag{5.17}$$

Tirando o mínimo múltiplo comum e eliminando dt, temos:

$$\frac{\partial C(S,t)}{\partial S(t)}S(t)r - C(S,t)r = -\frac{\partial C}{\partial t} - \frac{\partial^2 C}{\partial S^2} \tag{5.18}$$

Rearranjando os termos:

$$\frac{\partial C}{\partial t} + \frac{1}{2}\frac{\partial^2 C}{\partial S^2}dt + rS\frac{\partial C}{\partial S} - rC = 0 \tag{5.19}$$

Onde a equação acima é a famosa equação diferencial de Black e Scholes. Em particular, para a opção de compra, as condições de contorno são:

$$\begin{cases} C(0,t) = 0 \\ \lim_{S\to\infty} C(S,t) \to S \\ C(S,T) = \max(S-K,0) \end{cases} \tag{5.20}$$

5.2 Precificação de opções para modelo de difusão de preços com trajetórias diferenciáveis

Vamos desenvolver agora um exemplo onde poderemos identificar o raciocínio necessário para a obtenção da fórmula de Black e Scholes. Através de algumas simplificações em relação às hipóteses originais de Black e Scholes, aplicaremos o conceito de hedging dinâmico para obtermos uma equação diferencial determinística. Estas simplificações permitirão a análise da precificação de derivativos sem a necessidade de recorrermos às complexas propriedades do cálculo estocástico.

Consideremos:

$C(S,t)$ uma função, derivável um número suficiente de vezes, que representa o preço do derivativo

$S(t)$, com $t \in [0,T]$, o preço do ativo-objeto que segue um processo estocástico cujas trajetórias são diferenciáveis

Observe que, neste exemplo, $S(t)$ é diferenciável, ao contrário da formulação original de Black e Scholes onde o preço do ativo-objeto segue um movimento Browniano geométrico e que, portanto, não apresenta trajetórias diferenciáveis. Esta hipótese será essencial para que não seja necessária a utilização de conceitos de integrais ou diferenciais estocásticas.

Assim,

$$V_P = S(t) - \frac{1}{\frac{\partial C(S,t)}{\partial S(t)}} C(S,t) \tag{5.21}$$

Suponhamos agora um intervalo de tempo Δt suficientemente grande que cause uma alteração ΔV_P no valor da carteira. Teremos portanto:

$$\Delta V_P = \Delta S(t) - \frac{1}{\frac{\partial C(S,t)}{\partial S(t)}} \Delta C(S,t) \tag{5.22}$$

Vamos agora aproximar ΔV_P por dV_P, $\Delta S(t)$ por $dS(t)$ e $\Delta C(S,t)$ por $dC(S,t)$. Isto é possível pela propriedade de continuidade de V_P, S e C. Substituindo na equação (5.22), temos:

$$dV_P = dS(t) - \frac{1}{\frac{\partial (S,t)}{\partial S(t)}} dC(S,t) \tag{5.23}$$

A derivada total $dC(S,t)$ é dada por:

$$dC(S,t) = \frac{\partial C(S,t)}{\partial S} dS + \frac{\partial C(S,t)}{\partial t} dt \tag{5.24}$$

Note que nesta última passagem utilizamos as propriedades de cálculo determinístico. Uma vez que $S(t)$ segue um processo estocástico diferenciável, de acordo com nossas simplificações iniciais, podemos realizar cálculos baseados na integral de Riemann-Stieltjes, não sendo necessário recorrermos à integral estocástica.

Substituindo (5.24) em (5.23):

$$dV_P = dS(t) - \frac{1}{\frac{\partial C(S,t)}{\partial S(t)}} \left(\frac{\partial C(S,t)}{\partial S(t)} dS + \frac{\partial C(S,t)}{\partial t} dt \right) \tag{5.25}$$

Realizando a operação de multiplicação, temos:

$$dV_P = dS - dS - \frac{\frac{\partial C(S,t)}{\partial t}}{\frac{\partial C(S,t)}{\partial S(t)}} dt \tag{5.26}$$

que se reduz a

$$dV_P = -\frac{\frac{\partial C(S,t)}{\partial t}}{\frac{\partial C(S,t)}{\partial S(t)}} dt \tag{5.27}$$

Observe que os termos que refletem o comportamento aleatório representado por dS são eliminados. Com isso, a análise do valor da opção torna-se independente do caráter aleatório da flutuação de preço da ação em função do tempo. Neste caso, a aleatoriedade é eliminada através da forma com que é estabelecido o hedging dinâmico, isto é, a formação e o ajuste da carteira livre de risco.

Uma vez que a carteira dada pela equação (5.21) não apresenta risco, é de se esperar que, para não haver possibilidade de arbitragem, após um período de tempo Δt, a carteira tenha uma rentabilidade equivalente a uma aplicação em renda fixa valorizada a uma taxa livre de risco. Assim,

$$\Delta V_P = \left(S(t) - \frac{1}{\frac{\partial C(S,t)}{\partial S(t)}} C(S,t)\right) r \Delta t \tag{5.28}$$

Aproximando ΔV_P por dV_P e Δt por dt, temos:

$$dV_P = \left(S(t) - \frac{1}{\frac{\partial C(S,t)}{\partial S(t)}} C(S,t)\right) r dt \tag{5.29}$$

Igualando-se as equações (5.27) e (5.29), obtemos:

$$-\frac{\frac{\partial C(S,t)}{\partial t}}{\frac{\partial C(S,t)}{\partial S(t)}} dt = \left(S(t) - \frac{1}{\frac{\partial C(S,t)}{\partial S(t)}} C(S,t)\right) r dt \tag{5.30}$$

E eliminando dt:

$$-\frac{\frac{\partial C(S,t)}{\partial t}}{\frac{\partial C(S,t)}{\partial S(t)}} = \left(S(t) - \frac{1}{\frac{\partial C(S,t)}{\partial S(t)}} C(S,t)\right) r \tag{5.31}$$

Simplificando a equação acima, temos:

$$\frac{\partial C(S,t)}{\partial t} = rS(t)\frac{\partial C(S,t)}{\partial S(t)} - rC(S,t) \tag{5.32}$$

que pode ser rearranjada:

$$\frac{\partial C(S,t)}{\partial t} - rS(t)\frac{\partial C(S,t)}{\partial S(t)} + rC(S,t) = 0 \tag{5.33}$$

A equação acima representa uma equação diferencial parcial análoga à equação diferencial obtida por Black e Scholes vista na seção anterior. Se compararmos as equações (5.33) e (5.19), podemos perceber as alterações que a hipótese de diferenciabilidade de $S(t)$ causam na formulação original de Black e Scholes. Em ambos os casos, porém o componente aleatório desaparece e as equações diferenciais obtidas são determinísticas, ou seja, não-estocásticas.

Passemos agora à resolução da equação (5.33). De acordo com Foland (Foland,), uma forma de solução pode ser obtida pela introdução de uma variável auxiliar z, fazendo com que seja estabelecido um sistema de equações ordinárias. Assim, temos:

$$\begin{cases} \frac{dS}{dz} = rS \\ \frac{dt}{dz} = 1 \\ \frac{dC}{dz} = rC \end{cases} \tag{5.34}$$

onde consideramos S, t e C funções somente da variável z.

A condição inicial da equação é definida por:

$$(S,t,C)_{z=0} = (S_T, T, \Psi(S_T,T)) \tag{5.35}$$

onde S_T é um parâmetro que representa o valor da ação no vencimento e, para o caso de uma opção de compra com preço de exercício K no vencimento, como já visto anteriormente, a função Ψ é dada por:

$$\Psi(S_T, T) = \begin{cases} S_T - K \text{se } S_T \geq K \\ = 0 \quad \text{se } S_T < K \end{cases} \tag{5.36}$$

A solução deste sistema de equações, como pode ser verificado por substituição, é:

$$\begin{cases} S(z) = S_T \exp(rz) \\ t(z) = z + T \\ C(z) = \Psi(S_T, T)\exp(rz) \end{cases} \tag{5.37}$$

Expressando z e S_T como funções de S e t, utilizando as duas primeiras equações acima e substituindo na terceira, obtemos:

$$\begin{cases} z = t - T = -(T-t) \\ S_T = \frac{S(t)}{\exp[-r(T-t)]} = S(t)\exp[r(T-t)] \\ C(S,t) = \Psi(S\exp[r(T-t)]\exp[-r(T-t)] \end{cases} \tag{5.38}$$

Este exemplo, embora simples, apresenta propriedades interessantes. Nas condições propostas pelo exemplo, se $S(t)$, o preço do ativo-objeto, segue um processo aleatório com trajetórias diferenciáveis, então o preço da call antes do vencimento é simplesmente calculado de acordo com o seguinte procedimento:

a) Observa-se o valor da ação e corrige-se a valor futuro pela taxa livre de risco de acordo com o prazo para o vencimento $T - t$. Assim, $S_T = S\exp[r(T-t)]$

b) Compara-se este valor da ação corrigido com o preço de exercício da opção, encontrando-se o valor da opção no vencimento, ou seja, $\Psi(S_T, T)$. Caso S_T seja maior que o preço de exercício, então a opção possui prêmio equivalente ao valor intrínseco $S_T - K$. Caso contrário, a opção terá virado pó e portanto terá valor igual a 0

c) Traz-se $\Psi(S_T, T)$ a valor presente, pela taxa livre de risco, obtendo-se o valor da opção num momento qualquer t antes do vencimento. Ou seja: $C(S,t) = \exp[-r(T-t)]\Psi(S\exp[r(T-t)])$.

O valor da opção, neste caso, depende apenas do preço à vista da ação e do prazo para o vencimento, não sendo influenciado pelo caráter aleatório do preço da ação no tempo, nem por um parâmetro σ que caracteriza o grau de dispersão dos retornos no caso estudado por Black e Scholes. Esta fórmula parece ser mais adequada para avaliação de outros tipos de contratos de derivativos, como por exemplo, swaps, futuros ou termos. A suposição de diferenciabilidade parece portanto, não ser adequada para o modelamento de contratos de opções.

Devemos observar que o resultado obtido não depende da lei do processo $S.$, sendo igual para todos os processos cujas trajetórias são diferenciáveis. Vamos representar agora um processo específico $S.$ com trajetórias diferenciáveis. A característica deste processo é que suas distribuições no tempo $t = 0$, 1, 2, ..., T coincidem com as do movimento browniano padrão $B.$.

Definamos:

$$S_0 \equiv 0 \tag{5.39}$$

$$S_1 | S_0 \stackrel{\mathcal{D}}{=} S_0 \exp(X_1) \tag{5.40}$$

$$\vdots$$

$$S_{i+1}|S_i \stackrel{\mathcal{D}}{=} S_i \exp(X_i + 1) \tag{5.41}$$

com $X_1 = \mathcal{N}(0;1)$, ..., $X_{i+1} = \mathcal{N}(0;1)$

Ou seja, se $S_i = s$, então $S_{i+1} = s\exp(X_{i+1})$ onde $X_{i+1} \sim \mathcal{N}(0;1)$ e as variáveis $X_{k,s}, k = 1, 2, ..., T, s \in$ são independentes. Estes fatos determinam a distribuição de $S.$ nos instantes 0, 1, ..., T. Para definir $S.$ nos outros pontos do intervalo $[0, T]$, ligaremos os seus valores nos instantes 0, 1, ..., T através de funções suaves, garantindo a diferenciabilidade de suas trajetórias.

Seja então

$$f(x) = \sin\left(x\frac{\pi}{2}\right) \tag{5.42}$$

com $x \in [0, 1]$.

Observe que $f(x)$ está definida em $[0, 1]$, com $f(0) = 0$ e $f(1) = 1$ e que $f(\cdot)$ é infinitamente derivável. Além disso, a derivada à direita em $x = 0$ e a derivada à esquerda em $x = 1$ são iguais à 0. Vamos definir então, para $t \in [i, i+1]$:

$$S_t = S_i + f(t-i)(S_{i+1} - S_i) \tag{5.43}$$

A equação acima significa que ligaremos os pontos (i, S_i) e $(i+1, S_{i+1})$ pela função $f(\cdot)$ esticando-a apropriadamente. Das propriedades de $f(\cdot)$ apresentadas acima, segue que $S.$ tem trajetórias diferenciáveis.

Vamos supor agora que observaremos o processo $S.$ durante um intervalo de tempo $[i, i+\epsilon]$, onde $\epsilon > 0$ pode ser arbitrariamente pequeno. Das observações acima, podemos concluir, de maneira única, todas as derivadas da trajetória de $S.$ no ponto i. Portanto, saberemos o valor de $S.$ nos tempos $[i+\epsilon, i+1]$. Ou seja, quando a função que liga os pontos aleatórios de $S.$ é conhecida, como é o caso da função $f(\cdot)$ construída, então existe uma certa previsibilidade dos valores futuros, o que não acontece quando $S.$ é um movimento browniano geométrico. Este fato indica que os preços das opções devem ser distintos nos casos em que $S.$ tem trajetórias diferenciáveis e $S.$ tem um componente de movimento browniano. Adicionalmente, a previsibilidade de valores futuros sugere que o processo de trajetórias diferenciáveis não seja adequado para o modelamento do comportamento do preço da ação.

5.3 Aplicação do argumento de Black-Scholes para o processo de mistura

Nesta seção, vamos verificar a aplicabilidade do argumento de Black e Scholes para o processo de mistura definido no Capítulo 4.

Suponhamos que exista uma função

$$C(S,t), \qquad S \in \mathbb{R}^+,\ t \in [0, T] \tag{5.44}$$

que fornece o valor da opção para o tempo t e para o preço S da ação. Consideremos que esta função $C(\cdot,\cdot)$ é pelo menos duas vezes diferenciável em relação ao seu primeiro argumento S, e pelo menos uma vez diferenciável em relação ao seu segundo argumento t.

Vamos começar nossa análise a partir da fórmula do valor da carteira livre de risco. Temos, considerando uma posição comprada de uma ação e uma posição vendida de $\left(\frac{\partial C}{\partial S}\right)^{-1}$ opções:

$$V_P = S - C\left(\frac{\partial C}{\partial S}\right)^{-1} \tag{5.45}$$

Temos então

$$\mathrm{d}V_P = \mathrm{d}S + \mathrm{d}C\left(\frac{\partial C}{\partial S}\right)^{-1} \tag{5.46}$$

que segue da definição do valor da carteira V_P e do fato de que a diferencial estocástica da somatória é a somatória das diferenciais estocásticas. Cabe ressaltarmos que na equação (5.46), $\mathrm{d}S$ é diferente do conceito considerado na formulação original de Black e Scholes, tendo em vista que os processos de difusão do preço do ativo-objeto são diferentes.

A diferencial estocástica de C pode ser calculada pelo lema de Itô:

$$\mathrm{d}C = \frac{\partial C}{\partial S}\mathrm{d}S + \frac{1}{2}\frac{\partial^2 C}{\partial S^2}(\mathrm{d}S)^2 + \frac{\partial C}{\partial t}dt \tag{5.47}$$

onde $(\mathrm{d}S)^2$ é uma expressão cujo valor pode ser calculado através das seguintes regras [1]:

$$dt \times dt = dt \times \mathrm{d}B_t = \mathrm{d}B_t \times dt = 0, \qquad \mathrm{d}B_t \times \mathrm{d}B_t = dt \tag{5.48}$$

Poderíamos também expressar a diferencial estocástica de S em termos da diferencial estocástica do movimento browniano, mas prefiremos deixá-lo na forma de $\mathrm{d}S$. A vantagem desta representação está na facilidade de visualização do desaparecimento de $\mathrm{d}S$ que realizaremos no próximo passo do nosso raciocínio. Colocando (5.47) em (5.46), obtemos

$$\mathrm{d}P = -\left(\frac{\partial C}{\partial S}\right)^{-1}\left(\frac{1}{2}\frac{\partial^2 C}{\partial S^2}(\mathrm{d}S)^2 + \frac{\partial C}{\partial t}dt\right) \tag{5.49}$$

Observemos que na passagem de (5.46) para (5.49) o termo $\mathrm{d}S$ desapareceu, configurando no que se chama na literatura de "eliminação da aleatoridade". Esta eliminação da aleatoriedade é simplesmente decorrente do fato de que na expressão de $\mathrm{d}C$, $\mathrm{d}S$ está sendo multiplicado por $\frac{\partial C}{\partial S}$, enquanto $\mathrm{d}C$ na expressão de $\mathrm{d}V_P$ está sendo multiplicado por $(\frac{\partial C}{\partial S})^{-1}$.

Pelo resultado (5.49), temos que

$$rdt \times \left[S - C(S,t)\left(\frac{\partial C}{\partial S}\right)^{-1}\right] \tag{5.50}$$

Portanto, igualando (5.49) com (5.50) temos a seguinte equação:

$$-\left(\frac{\partial C}{\partial S}\right)^{-1}\left(\frac{1}{2}\frac{\partial^2 C}{\partial S^2}(\mathrm{d}S)^2 + \frac{\partial C}{\partial t}dt\right) = rdt \times \left[S - C(\frac{\partial C}{\partial S})^{-1}\right] \tag{5.51}$$

[1] veja Oksendal

Para continuarmos a análise de (5.51), precisamos calcular $\mathrm{d}S_t$. Para isto, podemos aplicar as regras de cálculo estocástico. Primeiro, aproveitaremos o fato de que a diferencial de soma é a soma das diferenciais:

$$\mathrm{d}S_t = S_0\Big\{\sum_{i=1}^{m} \mathrm{d}(\eta_i \exp\{\mu_i t + \sigma W_t\})\Big\} \tag{5.52}$$

A diferencial estocástica de cada membro pode ser calculado aplicando-se o lema de Itô:

$$\begin{aligned} \mathrm{d}(\eta_i \exp\{\mu_i t + \sigma W_t\}) &= \eta_i \cdot \Big([\mu_i \exp\{\mu_i t + \sigma W_t\} + \frac{1}{2}\sigma^2 \exp\{\mu_i t + \sigma W_t\}]\, dt \\ &+ \sigma \exp\{\mu_i t + \sigma W_t\}\, \mathrm{d}W_t\Big) \\ &= \eta_i \cdot \exp\{\mu_i t + \sigma W_t\} \cdot \Big(\Big[\mu_i + \frac{1}{2}\sigma^2\Big] dt + \sigma \mathrm{d}W_t\Big) \end{aligned} \tag{5.53}$$

Agora aplicando as regras dadas em (5.48), podemos concluir que

$$(\mathrm{d}S_t)^2 = S_0^2\Big[\sum_{i=1}^{m} \eta_i \cdot \exp\{\mu_i t + \sigma W_t\} \cdot \Big(\Big[\mu_i + \frac{1}{2}\sigma^2\Big] \times dt + \sigma \mathrm{d}W_t\Big)\Big]^2 \tag{5.54}$$

$$= S_0^2\Big[\sum_{i=1}^{m} \eta_i \cdot \exp\{\mu_i t + \sigma W_t\}\Big]^2 \sigma^2\, dt \tag{5.55}$$

$$= S_t^2 \sigma^2\, dt \tag{5.56}$$

Substituindo a expressão para $(\mathrm{d}S)^2$ obtida acima na equação (5.51) e dividindo ambos os membros por dt, obtemos a seguinte equação diferencial parcial

$$\frac{\partial C}{\partial t} = rC - rS\frac{\partial C}{\partial S} - \frac{1}{2}\frac{\partial^2 C}{\partial S^2}\sigma^2 S^2 \tag{5.57}$$

que é idêntica a equação obtida por Black e Scholes. Portanto, a sua solução ~~será~~.

$$\begin{aligned} C(S,t) &= S\Phi(d_1) - Ke^{r(t-T)}\Phi(d_2) \\ \text{onde } d_1 &= \frac{\log S/K + (r + \frac{1}{2}\sigma^2)\cdot(T-t)}{\sigma\sqrt{T-t}} \\ d_2 &= \frac{\log S/K + (r - \frac{1}{2}\sigma^2)\cdot(T-t)}{\sigma\sqrt{T-t}} \end{aligned} \tag{5.58}$$

Observe que a equação diferencial (5.57) coincidiu com a equação de Black e Scholes (5.19) obtida para um processo de modelamento do preço da ação baseado em um único browniano, graças à passagem da equação (5.54) à equação (5.55). Note que a coincidência não valeria no caso em que

$$S_t := S_0\big(\eta_1 \exp\{\mu_1 t + \sigma_1 W_t\} + \ldots + \eta_m \exp\{\mu_m t + \sigma_m W_t\}\big) \tag{5.59}$$

e pelo menos dois dos $\sigma_1, \ldots, \sigma_m$ sejam diferentes, pois partindo-se de:

$$\mathrm{d}S_t = S_0\Big\{\sum_{i=1}^{m} \mathrm{d}(\eta_i \exp\{\mu_i t + \sigma_i W_t\})\Big\} \tag{5.60}$$

temos

$$\begin{aligned}
\mathrm{d}(\eta_i \exp\{\mu_i t + \sigma_i W_t\}) &= \eta_i \cdot \Big(\Big[\mu_i \exp\{\mu_i t + \sigma_i W_t\} + \frac{1}{2}\sigma^2 \exp\{\mu_i t + \sigma_i W_t\}\Big] dt \\
&+ \sigma_i \exp\{\mu_i t + \sigma_i W_t\}\, \mathrm{d}W_t \Big) \qquad (5.61) \\
&= \eta_i \cdot \exp\{\mu_i t + \sigma_i W_t\} \cdot \Big(\Big[\mu_i + \frac{1}{2}\sigma_i^2\Big] dt + \sigma_i \mathrm{d}W_t \Big)
\end{aligned}$$

e portanto;

$$(\mathrm{d}S_t)^2 = S_0^2 \Big[\sum_{i=1}^{m} \eta_i \cdot \exp(\mu_i t + \sigma_i W_t) \sigma_i \mathrm{d}W_t \Big]^2 \qquad (5.62)$$

Observemos que neste caso, quando as variâncias não são iguais, a equação (5.51) não se transforma na equação diferencial parcial dada em (5.67). Portanto, sua solução dependerá do processo $W_\cdot$. Como veremos na próxima seção, este fato implica que a solução da equação não possa ser considerada o preço correto da opção.

Note que a equação de $(dS_t)^2$, quando pelo menos dois σ's são diferentes, contém termos estocásticos, dependentes de W_t. Por causa destes termos estocásticos, a solução, se existir, é uma função que depende de ω. Isto contradiz a nossa suposição inicial de que C depende somente de S e t. Portanto, embora a equação formalmente possa ter uma solução, a mesma não pode ser considerada como o valor da opção, pois não satisfaz a suposição inicial. Portanto, temos que modificar a suposição sobre C, incluindo sua dependência de ω. Mas tal dependência impossibilita a aplicação do lema de Itô. Chegamos assim, à seguinte conclusão: o raciocínio de estruturação de um portfólio livre de risco com ações e opções não se sustenta quando aplicado ao caso (5.59).

5.4 Considerações sobre o raciocínio de Black e Scholes

Nesta seção, resumiremos os procedimentos explanados nas seções anteriores para possibilitar uma melhor estruturação do raciocínio empregado por Black e Scholes para obtenção de soluções para o preço de derivativos. Assim, primeiramente, identificaremos as principais etapas ou passos do raciocínio e posteriormente, mostraremos como estes passos foram executados por Black e Scholes.

5.4.1 Etapas do raciocínio

Quando a questão é a obtenção do preço de um derivativo financeiro, um dos caminhos para sua solução é a utilização do famoso raciocínio de Black e Scholes, conforme estabelecido nas seções anteriores. Em linhas gerais, podemos extrair que os principais passos deste raciocínio são os seguintes:

P1 Assumir que o preço do derivativo é uma função C de certos argumentos que possuem certas propriedades.

P2 Construir um portfólio livre de risco cuja composição e manutenção utiliza a função C e quaisquer outros valores do mercado, que podem ser observados.

P3 Argumentar que o resultado decorrente da manutenção do portfólio livre de risco durante um pequeno intervalo de tempo deve ser equivalente ao ganho de uma aplicação , para este mesmo intervalo de tempo, de todo o valor do portfólio em um título de renda fixa, livre de risco. Este argumento implica a obtenção de uma equação envolvendo C.

P4 Achar a solução C da equação obtida pela aplicação do argumento do item **P3**.

P5 Verificar se a solução obtida em **P4** realmente satisfaz todas as condições impostas no item **P1**. Caso afirmativo, a solução C realmente representa o valor do derivativo.

O fato que queremos destacar no presente artigo é que existem trabalhos que implementam os passos **P1** - **P4** sem a preocupação de verificar a validade de **P5**, o que faz com que a solução encontrada C possa não ser considerada como o preço correto do derivativo. Na Subseção 5.5.1 apresentaremos um exemplo de uso incorreto do raciocínio de Black e Scholes. Na Subseção 5.5.2 mostraremos porque a não-validação de **P5** pode comprometer todo raciocínio. Antes de partirmos para esta apresentação, reveremos na Subseção 5.4.2, a execução de **P1-P5** para a precificação de uma opção de compra européia, segundo o trabalho original de Black e Scholes, para que possamos posteriormente discutir seu uso inadequado. Na nossa opinião, a utilização inapropriada em alguns estudos do raciocínio descrito pelos passos acima é fruto da intenção de se repetir literalmente a execução dos passos **P1-P5** do trabalho clássico [2] para situações que são bem distintas da situação abordada por Black e Scholes. Finalmente, na Subseção 5.5.3 discutiremos uma alternativa ao raciocínio referentes aos passos **P1-P5**, que fundamenta-se na abordagem através de martingais.

5.4.2 Precificação de opção européia

Nesta seção, vamos mostrar como os passos **P1-P5** apresentados anteriormente são executados no trabalho de Black e Scholes, para a precificação de uma opção de compra do tipo europeu.

Denotaremos por S_t o preço de uma ação no tempo t. Designaremos por $W_t, t \in [0, \infty)$ o movimento browniano padrão. Vamos também supor que existam duas constantes μ e σ tais que

$$S_t = \exp\{\mu\, t + \sigma\, W_t\}, \qquad t \in [0, \infty)$$

Vamos investigar a seguinte pergunta: qual o valor no tempo $t = 0$ de um contrato de opção de compra com o vencimento no tempo $t = T$ com preço de exercício K. Observaremos, que [BS] não só responde a esta pergunta mas também possibilita a identificação do valor deste contrato em qualquer instante t entre 0 e T. A resposta é consequência da execução dos itens **P1-P5** que designaremos por **E1-E5** respectivamente.

E1 Vamos assumir que se no tempo t o preço de ação é S_t, então preço de opção será dado por $C(S_t, t)$, onde C é uma função que depende de duas variáveis S_t e t somente. Além disso, assumiremos que C é pelo menos diferenciável duas vezes em relação à variável S_t e pelo menos diferenciável uma vez em relação à variável t.

[2]veja Black e Scholes

E2 Vamos designar por P o portfólio composto de acordo com a seguinte regra: no tempo t e com o valor S_t da ação neste tempo, o portfólio contém uma ação comprada e $(\partial C(S_t,t)/\partial S_t)^{-1}$ opções vendidas. Notaremos que, por **E1**, o valor deste portfólio depende efetivamente de somente dois argumentos, que são S_t e t. Assim, o valor do portfólio é dado por:

$$P(S_t,t) = S_t - \left(\frac{\partial C}{\partial S}\right)^{-1} C(S_t,t) \tag{5.63}$$

onde $\partial C/\partial S$ deve ser entendido como a derivada parcial de C em relação de S_t avaliada no ponto (S_t,t).

Segue-se diretamente de (5.63) e **E1** que

$$\frac{\partial P(S_t,t)}{\partial S_t} = 0 \tag{5.64}$$

A igualdade (5.64) pode ser interpretada como a imunidade do portfólio P em relação ao risco de alteração do preço do ativo. De fato, suponha que S_t sofreu uma pequena mudança de valor dS. Então $C(S_t,t)$ teve seu valor alterado em $(\partial C(S_t,t)/\partial S_t)dS$ adicionado de um valor infinitesimal referente a dS. Portanto, a variação absoluta do portfólio dado por

$$P(S_t+dS,t) - P(S_t,t) \tag{5.65}$$

é infinitesimal em dS. Daí, para concluirmos (5.64), ou seja, que o valor da carteira não sofre alterações com variações infinitesimais do preço do ativo-objeto, só falta então tomar o limite $dS \to 0$, e para concluirmos sobre a imunidade ao risco, podemos utilizar o argumento de que esta parte infinitesimal desaparece se o ajuste do portfólio é realizado de maneira contínua.

Podemos notar que existe uma certa discrepância entre (5.64) e a sua interpretação. Esta discrepância está induzida pelo simples fato de que o valor da ação não pode mudar sozinho, mas a mudança de S_t só ocorre com a mudança do tempo. Por isso, a expressão correta da variação de portfólio é

$$P(S_t+dS,t+dt) - P(S_t,t) \tag{5.66}$$

Observaremos que a tentativa de considerarmos (5.66) ao invés de (5.65) implica a passagem para diferencial completo de P. Este diferencial completo deve ser um diferencial estocástico, pois S é uma função de um browniano. Porém, isto seria incoerente como resto do argumento, tendo em vista que este diferencial está sendo usado na execução dada em **E3** para expressar o resultado referente à manutenção do portfólio P. Para justificar que (5.65) é mais apropriado de que (5.66) usa-se geralmente o argumento que afirma que a contribuição de dt à variação de P é infinitesimal em comparação com a contribuição de dS.

E3 O resultado líquido durante um pequeno intervalo de tempo $[t,t+dt]$ devido à manutenção do portfólio P expressa-se por $\mathrm{d}P(S_t,t)$, onde d deve ser entendido como a diferencial estocástica, para ser distingida de d, que usamos para designar a diferencial usual. Não pretendemos discutir neste trabalho a relevância de $\mathrm{d}P(S_t,t)$ para sua interpretação. Assim, considerando o argumento em **P3**, temos a seguinte equação:

$$\frac{\partial C}{\partial t} = rC - rS\frac{\partial C}{\partial S} - \frac{1}{2}\frac{\partial^2 C}{\partial S^2}\sigma^2 S^2 \tag{5.67}$$

E4 A equação (5.67) é uma equação diferencial parcial, que pode ser conduzida à equação de difusão do calor através de transformações apropriadas de variáveis. A solução desta equação que, satisfaz as condições de contorno que expressam o preço de uma opção de compra no vencimento, ou seja, $C(S_T, T) = \max[S_T - K, 0]$, é dada por:

$$\begin{aligned} C(S,t) &= S\Phi(d_1) - Ke^{r(t-T)}\Phi(d_2) \qquad (5.68)\\ \text{onde } d_{1,2} &= \frac{\log S/K + (r \pm \frac{1}{2}\sigma^2)\cdot(T-t)}{\sigma\sqrt{T-t}}\\ \text{e } \Phi(x) &= \frac{1}{\sqrt{2\pi}\sigma}\int_{-\infty}^{x} \exp\{-\frac{t^2}{2\sigma^2}\}dt \end{aligned}$$

E5 O fato de a solução da equação diferencial (5.67) satisfazer as condições assumidas em **E1** segue da forma da própria equação (5.67) ou diretamente da sua expressão explícita (5.68).

5.5 Restrição à aplicabilidade do raciocínio de Black e Scholes

5.5.1 Um exemplo de uso incorreto do argumento de Black e Scholes

Uma execução incorreta dos passos **P1-P5** será apresentada nesta seção. Trataremos do caso de uma precificação de um derivativo num exemplo particular, que construímos com o objetivo de propiciar uma maior transparência na exibição dos erros da execução. Embora artificial, este exemplo tem muitos fatores em comum com os modelos que vimos na literatura e cujas soluções pelo método de Black e Scholes inspiraram o presente artigo.

O problema estabelecido é encontrar o preço de uma concessão de exploração de uma fazenda de produção de leite, que será descrito pelo seguinte modelo matemático. Designaremos por M_t o número de cabeças de vacas desta fazenda no tempo t, e por S_t o preço de venda de uma unidade, por exemplo, um litro de leite no tempo t. Assumiremos que durante um intervalo pequeno de tempo $[t, t+dt]$ os dividendos ou rendimentos obtidos através da venda de leite seja $k \cdot M_t \cdot S_t dt$, onde k é uma certa constante, cujo valor pode depender da taxa de produção de leite por vaca, das despesas de manutenção da fazenda e dos impostos. Observaremos que em modelos de exploração de outros recursos, como por exemplo, madeira ou petróleo, os rendimentos correspondentes podem ter uma estrutura muito mais complexa, em particular, devido à possibilidade de existência, nestes modelos de opções embutidas de suspenção e/ou diminuição temporária da exploração. Suponhamos que $S_t, t \geq 0$, e $M_t, t \geq 0$, são movimentos brownianos geométricos dados por:

$$S_t = S_0 \exp\{\mu t + \sigma W_t\}, \qquad M_t = M_0 \exp\{\nu t + \kappa B_t\} \qquad (5.69)$$

onde $W_t, t \geq 0$, e $B_t, t \geq 0$, são dois movimentos brownianos padrão independentes e μ, σ, ν, κ são constantes. Vamos considerar também que o ativo leite é um produto estocável. Essa hipótese é essencial para que possamos compor um portfólio livre de risco. A seguir, em **R1-R5**, vamos repetir os passos **P1-P5**, para o modelo acima descrito como objetivo de precificarmos a concessão da exploração da fazenda.

R1 Designaremos por C o valor de concessão da fazenda. Suponhamos que, para qualquer que seja t, C depende somente de t, de S_t e de M_t, ou seja, $C = C(S_t, M_t, t)$. Vamos considerar tembém que esta função é diferenciável pelo menos duas vezes em relação a S_t e a M_t, e pelo menos diferenciável uma vez em relação a t.

R2 Denotaremos por P um portfólio composto de um direito de explorar a fazenda e de uma obrigação de vender $\partial C(S_t, M_t, t)/\partial S_t$ unidades de leite. Com isso, o valor de portfólio no tempo t é

$$P(S_t, M_t, t) = C(S_t, M_t, t) - \frac{\partial C(S_t, M_t, t)}{\partial S_t} S_t \tag{5.70}$$

Usando um argumento idêntico àquele usado em **E2**, temos que

$$\frac{\partial P}{\partial S} = 0 \tag{5.71}$$

Onde a interpretação da equação acima seria que P é livre de risco. Aceitaremos por enquanto essa interpretação, embora posteriormente mostraremos que ela é errada.

R3 O rendimento do portador do portfólio P dado em (5.70) durante um intervalo de tempo dt tem os componentes: *(1)* os ganhos provenientes da venda do leite kS_tM_tdt. *(2)* As despesas para a manutenção da posição vendida ou curta, que existem em decorrência de custos de oportunidade. Assumindo que tais custos são constantes e denominando-os por λ, temos a seguinte expressão para estas despesas: $\lambda S_t(\partial C/\partial S)dt$.

Assim, o resultado líquido da carteira, seguindo a discussão apresentada em **E3** deve ser representado pelo diferencial completo de P no ponto (S_t, M_t, t). Com isso, temos:

$$\mathrm{d}P(S_t, M_t, t) = \mathrm{d}(C(S_t, M_t, t)) - \frac{\partial C(S_t, M_t, t)}{\partial S_t}\mathrm{d}S_t \tag{5.72}$$

$$= \frac{\partial C}{\partial t}dt + \frac{1}{2}\frac{\partial^2 C}{\partial S^2}(\mathrm{d}S_t)^2 + \frac{\partial C}{\partial M}\mathrm{d}M_t + \frac{1}{2}\frac{\partial^2 C}{\partial M^2}(\mathrm{d}M_t)^2 \tag{5.73}$$

onde a nomenclatura de $(\mathrm{d}S_t)^2$ e de $(\mathrm{d}M_t)^2$ segue um formalismo teórico que o leitor pode encontrar em Oksendal[3]. Observe que na passagem de (5.72) para (5.73) usamos o lema de Itô, pois os processos S e M dependem dos movimentos brownianos W_t e B_t. Usando o princípio apresentado em **P3** e calculando $\mathrm{d}S_t$ e $\mathrm{d}M_t$ a partir de (5.69), obtemos a seguinte equação diferencial estocástica:

$$\begin{aligned}\left(C - \frac{\partial C}{\partial S}S_t\right)dt &= \frac{\partial C}{\partial t}dt + \frac{\sigma^2}{2}\frac{\partial^2 C}{\partial S^2}S_t^2 dt + \left(\nu + \frac{\kappa^2}{2}\right)\frac{\partial C}{\partial M}M_t dt \\ &+ \kappa\frac{\partial C}{\partial M}M_t\mathrm{d}B_t + \frac{\kappa^2}{2}\frac{\partial^2 C}{\partial M^2}M_t^2 dt + kS_tM_tdt - \lambda S_t\frac{\partial C}{\partial S}dt\end{aligned} \tag{5.74}$$

[3]veja Oksendal

R4 É conhecido, pela teoria de equações diferenciais estocásticas, que a solução da equação (5.74) existe e é única, desde que as condições de contorno sejam apropriadamente determinadas. De acordo com as considerações financeiras e de bom senso podemos estabelecer no nosso modelo todas as condições necessárias para garantir a existência e unicidade de tal solução. Vamos ressaltar que não conhecemos a expressão analítica da solução de (5.74), mas este desconhecimento não afetará o nosso argumento de que existe uma utilização inapropriada na execução indiscriminada dos passos **P1-P5**.

R5 Também da teoria das equações diferenciais estocásticas, sabe-se que a solução de (5.74) será pelo menos diferenciável duas vezes em relação a M_t e a S_t e pelo menos diferenciável uma vez em relação a t. Porém, tembém é conhecido que o valor desta solução no tempo t pode depender, e no caso discutido aqui realmente depende, da porção da trajetória do movimento browniano B entre os instantes 0 e t. Designando esta porção por $B_{[0,t]}$, temos então que $C = C(S_t, M_t, B_{[0,t]}, t)$.

Vamos notar que a conclusão de **R5** contradiz a hipótese sobre C assumida em **R1**, onde estabelecemos $C = C(S_t, M_t, t)$. Isto nos fornece uma justificativa formal de afirmar que então todo raciocínio **R1-R5** é incorreto, conforme mostraremos na próxima seção .

5.5.2 Discussão do uso incorreto do argumento de Black e Scholes

Vamos assumir que a solução C obtida em (5.74) realmente depende somente de S, M e t. Afirmamos, com esta suposição, que executando **R1-R5** não obtemos o valor justo da concessão. O argumento é o seguinte. A imunidade do portfólio P em relação ao risco foi concluída em **R2** repetindo o argumento apresentado em **E2**. Mas no caso **R**, em comparação com **E**, o portfólio possui mais um argumento, dado por M_t. Agora lembremos que S_t pode mudar somente com a mudança de tempo. Em **E2** a contribuição do acréscimo de t na variação do portfólio foi ignorada. Em **R2** porém, este acréscimo de t induz uma variação dM_t em M_t. O valor da contribuição de dM_t na mudança do valor do portfólio pode ser da mesma ordem de grandeza da contribuição ocasionada por dS_t. Se isto for verdade, então o portfólio P^* composto por

$$P_t^* = C(S, M, t) - \frac{\partial C}{\partial S} S_t - \frac{\partial C}{\partial M} M_t \tag{5.75}$$

é imune ao risco e P não será. Pode-se discutir um problema com a equação (5.75), porque a medida da terceira parcela é dada em "cabeças de vacas", enquanto as medidas de todas as outras parcelas estão estabelecidas em "dinheiro". Uma possível solução para o problema seria simplesmente introduzir no nosso modelo o processo $U_t, t \geq 0$, que expressaria o valor de uma cabeça de vaca no tempo t. Com isso a expressão correta de P^* seria ajustada através da multiplicação da terceira parcela por U_t. Mas este procedimento ainda não possibilita fechar o argumento **R1-R5**, pois a introdução de mais uma variável, representada pelo processo U_t, suscita a dúvida se C é uma função que também depende desta variável. Neste caso, **R1-R5** deve ser repetido com a suposição $C = C(S_t, M_t, U_t, t)$. Não prosseguiremos com esta análise, uma vez que ela depende de cada caso a ser considerado. A nossa intenção era somente mostrar os pontos críticos do argumento de Black e Scholes que exigem um particular cuidado quando generalizado para casos mais complexos.

A correção de P para P^* demonstrada acima é bem óbvia, embora vimos exemplos na literatura, onde não se cogita nem a consideração de P^*. Ou seja, o raciocínio de Black e Scholes é muitas vezes utilizado indiscriminadamente, podendo levar à desconsideração de fatores que podem ser relevantes como é o caso da parcela $-\left(\frac{\partial C}{\partial M}M_t\right)$ da equação (5.75). Achamos que uma das razões desta negligência é a tentativa de seguir literalmente as idéias expostas por Dixit e Pindyck[4] que, na maioria dos seus exemplos, consideram uma opção como a função de um só ativo.

Vamos confessar não temos uma receita universal para verificar se P e melhor que P^* ou vice-versa. Segue da dicussão acima que esta decisão depende da comparação entre as ordens de grandezas das contribuições de dS_t e dM_t no valor do portfólio P.

Porém, o fato de que a dependência de C do B, que não foi assumida em **R1**, invalida todo o raciocínio estabelecido em **R1-R4**, deve estar claro para leitor. De fato, mostramos acima que todos os argumentos de C devem ser tomados em conta para a composição do portfólio livre de risco. Mas $B_{[0,t]}$ foi ignorado, tanto em P quanto em P^*. O problema com B é que é muito difícel comparar a contribuição do seu "acréscimo" à variação do valor do portfólio com as correspondentes contribuições das variações de dS_t e dM_t. Isto acontece, em particular, em casos em que não é possível encontrar uma expressão analítica da equação diferencial estocástica. Sendo que tais casos são muito frequentes,[5] o erro da ignorância de **P5** também é frequente. Infelizmente, não podemos sugerir uma solução completa e eficiente para estes casos. Um caminho alternativo seria a utilização da abordagem através de martingais, que discutiremos na Seção V, a seguir.

5.5.3 Comparação com a abordagem através de martingais

Sabe-se que a precificação de derivativos pode ser feita pela através da utilização de martingais[6]. É natural, portanto, investigarmos se esta abordagem funcionará nos casos para os quais o raciocínio **P1-P5** não se aplica no sentido que mostramos acima. Começaremos com um exemplo onde realmente a abordagem através de martingais funciona. Depois indicaremos os obstáculos para a utilização de martingais para os modelos que vimos na literatura.

Consideraremos o problema de precificação de opção de uma compra européia para o seguinte modelo da dinâmica de preço do ativo-objeto. Existem n processos $S_t^k, t \geq 0$, definidos por

$$S_t^k = S_0 \exp\{\mu_k t + \sigma_k W_t\}, k = 1, 2, \dots, n \tag{5.76}$$

onde $\mu_1, \dots, \mu_n, \sigma_1, \dots, \sigma_n$ são constantes. No instante inicial $t = 0$, um desses processos está sendo escolhido com as relativas probabilidades $p_1, p_2, \dots, p_n$, e o preço da ação $S_t, t \geq 0$, segue o processo escolhido para todo $t \geq 0$.[7] É fácil ver que se pelo menos dois dos $\sigma_1, \dots, \sigma_n$ são distintos, então e equação obtida pelo argumento **E1-E3** é uma equação diferencial estocástica. Portanto, conforme apresentado na Seção IV, uma análise detalhada é necessária antes de poder afirmar que a solução desta equação fornece o preço correto da opção.[8] Por outro lado, através de martingais conclui imediatamente

[4]veja Dixit e Pindyck

[5]O exemplo da Seção V mostra que um pequeno ajuste no modelo original de Black e Scholes já faz com que a equação diferencial transforme-se em uma equação diferencial estocástica implicando a dependência direta de C por W.

[6]veja Neftci

[7]Num modelo mais próximo à realidade, o preço da ação poderia seguir processos distintos nos intervalos de tempo distintos, com a distribuição de instantes de mudança conforme um processo pontual que pode ser, por exemplo, um processo de Poisson. Esse processo é, porém, muito complicado para as necessidades da nossa exibição.

[8]O leitor poderá ver que a abordagem através de martingais resolve o problema de precificação da opçãosó quando combinada com outras idéias. Isto pode indicar para o leitor que o problema é complexo demais para poder ser resolvido somente seguindo-se **E1-E4**.

que

$$C(S_t,t) = C_k(S_t^k,t) \tag{5.77}$$

se o preço da ação segue o processo $S_t^k, t \geq 0$ onde $C_k(S_t^k,t)$ está dado por (5.68) com $S = S_t^k$ e $\sigma = \sigma_k$. Argumentando rigorosamente, a fórmula (5.77) ainda não resolve o problema, porque para obtermos o valor da opção no tempo t precisamos saber qual processo, dentre os n possíveis, foi escolhido para ser seguido. Isto porém, pode ser resolvido utilizando uma idéia do raciocínio **E1-E5**. Lembraremos que durante a explanação de **E1-E5** aceitamos a possibilidade de ajuste contínuo do portfólio P. Mas este ajuste é equivalente à possibilidade de observação contínua do processo $S_t, t \geq 0$. Estas observação fornecem uma amostra de tamanho arbitrariamente grande durante um intervalo de tempo $[0, dt]$ arbitrariamente pequeno. Aplicando à esta amostra qualquer método estatístico assintoticamente consistente, um investidor pode inferir corretamente os valores de σ_k e μ_k durante um intervalo arbitrariamente pequeno $[0, dt]$. As perdas do investidor durante este intervalo, enquanto só ele observa o processo, podem ser ignoradas. Portanto, (5.77) funcionará para $t > 0$. Assim, ficou para ser resolvido somente a questão do preço da opção no tempo $t = 0$. Aplicando o princípio de expectativas homogêneas dos participantes do mercado, temos que

$$C(S_0,0) = \sum_{k=1}^{n} p_k C_k(S_0,0) \tag{5.78}$$

Se o leitor ficou convencido de que a aplicação de raciocínio de Black e Scholes para os modelos parecidos com o da Seção III, é problemática, então a pergunta natural que surge é por que estes modelos não são tratados pela abordagem através de martingais. O exemplo acima fortalece esta dúvida. Uma possível resposta, porém, pode ser levantada. Os modelos utilizados são geralmente muito mais complexos do que o modelo da Seção III. Devido a esta complexidade, é extremamente difícil achar a transformação que faz com que o ativo-objeto, ou seu equivalente, seja um martingal. Sem essa transformação, a referida abordagem não permitirá a obtenção de uma solução. Por outro lado, a execução de **P1-P3** sempre dá uma equação que sempre pode ser resolvida, na pior hipótese, pelo um método numérico. Na maioria dos casos realmente obtém-se somente uma solução numérica, mas acredita-se que ela está suficientemente próxima à solução verdadeira. O fato, para o qual estamos apontando neste artigo, é que a solução verdadeira pode não expressar o preço correto, se **P5** não vale.

Apresentaremos agora brevemente um dos modelos complexos acima referidos. Ele será semelhante ao modelo da Seção III com as seguintes modificações. Existe um custo da manutenção da fazenda (limpeza, visitas de veterinários, fiscais da produção, etc.) que é um valor fixo X caso o leite esteja sendo vendido e é zero caso não haja venda de leite. Existe também a possibilidade de suspender temporariamente a venda de leite, caso os lucros de venda não cubram X. Por causa desta possibilidade, o espaço probabilístico onde B e W são definidos, está sendo partido em regiões, com estratégias diferentes a serem adotadas em cada uma das regiões. Neste caso e na maioria dos outros casos, não conhecemos as expressões analíticas das fronteiras dessas regiões. Este fato dificulta a procura da transformação deste espaço probabilístico que faz com que os produtos financeiros envolvidos sejam martingais.

Chapter 6

Resolução da equação diferencial de Black e Scholes

Na maioria dos livros sobre derivativos, afirma-se, sem demonstração, que a equação diferencial de Black e Scholes é equivalente à equação de difusão do calor da física. Neste capítulo, inicialmente faremos uma breve explanação sobre a equação de difusão do calor, citando suas principais características e apresentando sua solução, dada uma condição inicial. Posteriormente, mostraremos como transformar, através de sucessivas mudanças de variáveis, a equação diferencial de Black e Scholes na equação de difusão do calor. Nosso objetivo, portanto, é obter a fórmula final de Black e Scholes para avaliação de preços de opções de compra. Na primeira seção, apresentaremos, sem demonstrar, a solução da equação diferencial do calor que servirá de base para a resolução da equação diferencial de Black e Scholes. Na seção seguinte, aplicando o resultado da equação de difusão do calor, obteremos a fórmula de Black e Scholes para precificação de opções de compra. Na terceira seção, através da avaliação no mundo risco-neutro, isto é, supondo que todos os ativos têm retorno esperado igual à taxa de juros livre de risco, chegaremos à mesma fórmula de Black e Scholes obtida através da resolução da equação diferencial.

6.1 Equação de difusão do calor

A equação do calor ou a equação de difusão do calor é dada, na sua forma geral, por:

$$\frac{\partial u}{\partial \zeta} = \frac{\partial^2 u}{\partial x^2} \tag{6.1}$$

onde $u = u(x, \zeta)$ representa a temperatura de um ponto de uma barra de material uniforme, extremamente fina, com as extremidades perfeitamentre isoladas, ζ corresponde ao parâmetro tempo e x, à distância entre uma das extremidades da barra e o ponto analisado.

Encontrar uma solução $u(x, \zeta)$ para a equação diferencial acima significa obter uma fórmula que indique a temperatura, para um dado instante ζ, em um ponto que dista x de uma determinada extremidade. Analogamente, a fórmula de Black e Scholes fornece $C(S, t)$, ou seja, o preço da opção de compra em função do preço da ação $S(t)$ para qualquer instante t antes do vencimento.

A equação de difusão tem sido estudada por quase dois séculos como modelo de fluxo do calor em um meio contínuo, sendo um dos mais bem-sucedidos e utilizados modelos da matemática. Curiosamente, a fórmula de Black e Scholes obtida através do auxílio da solução da equação diferencial do calor, pode ser considerada também um dos modelos mais bem-sucedidos e aplicados no mundo das finanças.

Voltando ao estudo da equação diferencial do calor, podemos citar as seguintes propriedades [1]:

a) a equação é linear, ou seja, se u_i, $i = 1, ..., m$ são soluções da equação (6.1) e α_i são constantes quaisquer, então $\sum_{i=1}^{m} \alpha_i u_i$ também é solução da equação (6.1)

b) a equação é de segunda ordem, isto é, a maior ordem de derivada presente na equação é a segunda: $\partial^2 u/\partial x^2$

c) a equação é parabólica

d) a equação possui solução analítica

Vamos apresentar a solução analítica para a equação de difusão do calor, sujeita a condições específicas. Seja

$$\frac{\partial u}{\partial \zeta} = \frac{\partial^2 u}{\partial x^2} \tag{6.2}$$

sendo $-\infty < x < \infty$, $\zeta > 0$ com a condição inicial arbitrária $u(x,0) = u_0(x)$ e condição de contorno em $x = 0$ tal que $u(0,\zeta) = 1$.

Sob estas condições, a solução da equação (6.2), é dada por:

$$u(x,\zeta) = \frac{1}{2\sqrt{\pi\zeta}} \int_{-\infty}^{\infty} u_0(s) \exp\Big(-\frac{(x-s)^2}{4\zeta}\Big) ds \tag{6.3}$$

Cabe ressaltar que nesta seção apenas descrevemos a solução da equação (6.2). Utilizaremos este resultado para, na próxima seção, calcularmos detalhadamente a solução da equação diferencial de Black e Scholes. Maiores detalhes sobre a resolução da equação de difusão de calor podem ser encontrados em livros de cálculo onde a derivação da solução é baseada no delta de Dirac e na função de Heaviside ou em transformadas de Fourier.

6.2 Equação diferencial de Black e Scholes

Baseados na solução da equação de difusão do calor, vamos resolver a equação diferencial de Black e Scholes para opções de compra. Do capítulo anterior, temos a seguinte equação:

$$\frac{\partial C}{\partial t} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 C}{\partial S^2} + rS\frac{\partial C}{\partial S} - rC = 0 \tag{6.4}$$

com as seguintes condições de contorno:

$$\begin{cases} C(0,t) = 0 \\ \lim_{S\to\infty} C(S,t) \to S \\ C(S,T) = \max(S-K,0) \end{cases} \tag{6.5}$$

Através de algumas manipulações matemáticas, vamos transformar a equação diferencial de Black e Scholes na equação de difusão do calor dada por (6.2).

[1] veja Wilmott, Howison, Dewynne

Os passos para esta transformação não são imediatos. Primeiramente, vamos eliminar S e S^2 da equação (6.4) que multiplicam os termos $\frac{\partial C}{\partial S}$ e $\frac{\partial^2 C}{\partial S^2}$, respectivamente, através da troca de variáveis estabelecida abaixo:

$$\begin{cases} S = K \exp x \\ t = T - \frac{\zeta}{\frac{1}{2}\sigma^2} \\ C = Kv(x,t) \end{cases} \tag{6.6}$$

onde K é o preço de exercício e T o prazo para o vencimento.

Esta troca de variáveis induz as seguintes relações :

$$S = K \exp x \Longrightarrow \exp x = \frac{S}{K} \Longrightarrow x = \ln \frac{S}{K} \tag{6.7}$$

$$\frac{\partial x}{\partial S} = \frac{1}{\frac{S}{K}} \frac{1}{K} = \frac{1}{S} \tag{6.8}$$

$$\frac{\partial^2 x}{\partial S^2} = -\frac{1}{S^2} \tag{6.9}$$

$$t = T - \frac{\zeta}{\frac{1}{2}\sigma^2} \Longrightarrow -\frac{\zeta}{\frac{1}{2}\sigma^2} = t - T \Longrightarrow \zeta = \frac{1}{2}(T-t)\sigma^2 \tag{6.10}$$

$$\frac{\partial \zeta}{\partial t} = -\frac{1}{2}\sigma^2 \tag{6.11}$$

As derivadas parciais da equação (6.4) podem ser calculadas, aplicando conceitos de cálculo determinístico, uma vez que a equação (6.4) é determinística, pois a parte estocástica foi eliminada, como evidenciado no Capítulo 5. Assim, $\frac{\partial C}{\partial t}$ é dada por:

$$\frac{\partial C}{\partial t} = K \frac{\partial v}{\partial x}\frac{\partial x}{\partial t} + K \frac{\partial v}{\partial \zeta}\frac{\partial \zeta}{\partial t} \tag{6.12}$$

Sendo $\frac{\partial x}{\partial t} = 0$ e substituindo (6.11) em (6.12):

$$\frac{\partial C}{\partial t} = -\frac{1}{2}\sigma^2 K \frac{\partial v}{\partial \zeta} \tag{6.13}$$

Calculando $\frac{\partial C}{\partial S}$, obtemos:

$$\frac{\partial C}{\partial S} = K \frac{\partial v}{\partial x}\frac{\partial x}{\partial S} + K \frac{\partial v}{\partial \zeta}\frac{\partial \zeta}{\partial S} \tag{6.14}$$

Sendo $\frac{\partial \zeta}{\partial S} = 0$ e substituindo (6.8) em (6.14), temos:

$$\frac{\partial C}{\partial S} = \frac{1}{S^2} K \frac{\partial v}{\partial x}\frac{\partial x}{\partial S} \tag{6.15}$$

Para calcularmos $\frac{\partial^2 C}{\partial S^2}$, vamos fazer:

$$\frac{\partial^2 C}{\partial S^2} = \frac{\partial}{\partial S}\Big(\frac{\partial c}{\partial S}\Big) = \frac{\partial}{\partial S}\Big(\frac{K \partial v}{\partial S}\Big) = K \frac{\partial}{\partial S}\Big(\frac{\partial v}{\partial S}\Big) = K \frac{\partial}{\partial S}\Big(\frac{\partial v}{\partial x}\frac{\partial x}{\partial S} + \frac{\partial v}{\partial \zeta}\frac{\partial \zeta}{\partial S}\Big) \tag{6.16}$$

Sendo $\frac{\partial \zeta}{\partial S} = 0$ então:

$$\frac{\partial^2 C}{\partial S^2} = K\frac{\partial}{\partial S}\left(\frac{\partial v}{\partial x}\frac{\partial x}{\partial S}\right) \tag{6.17}$$

Aplicando a propriedade de diferenciação de produtos de funções dependentes de S, temos:

$$\begin{aligned} \frac{\partial^2 C}{\partial S^2} &= K\left[\frac{\partial}{\partial S}\left(\frac{\partial v}{\partial x}\right)\frac{\partial x}{\partial S} + \frac{\partial}{\partial S}\left(\frac{\partial x}{\partial S}\right)\frac{\partial v}{\partial x}\right] \\ &= K\left[\frac{\partial}{\partial x}\left(\frac{\partial v}{\partial x}\right)\frac{\partial x}{\partial S}\frac{\partial x}{\partial S} + \frac{\partial}{\partial \zeta}\left(\frac{\partial v}{\partial x}\right)\frac{\partial \zeta}{\partial S}\frac{\partial x}{\partial S} + \frac{\partial^2 x}{\partial S^2}\frac{\partial v}{\partial x}\right] \end{aligned} \tag{6.18}$$

Como $\frac{\partial \zeta}{\partial S} = 0$ então:

$$\frac{\partial^2 C}{\partial S^2} = K\left[\frac{\partial^2 v}{\partial x}\left(\frac{\partial x}{\partial S}\right)^2 + \frac{\partial^2 x}{\partial S^2}\frac{\partial v}{\partial x}\right] \tag{6.19}$$

Substituindo (6.18) e (6.19) na equação acima:

$$\frac{\partial^2 C}{\partial S^2} = K\left[\frac{\partial^2 v}{\partial x}\left(\frac{1}{S^2}\right) - \frac{1}{S^2}\frac{\partial v}{\partial x}\right] = \frac{1}{S^2}K\frac{\partial^2 v}{\partial x^2} - \frac{1}{S^2}K\frac{\partial v}{\partial x} \tag{6.20}$$

Assim, substituindo (6.13), (6.15) e (6.20) na equação básica (6.4) e levando em consideração (6.6):

$$-\frac{1}{2}\sigma^2 K\frac{\partial v}{\partial \zeta} + \frac{1}{2}\sigma^2 S^2\left(\frac{1}{S^2}K\frac{\partial^2 v}{\partial x^2} - \frac{1}{S^2}K\frac{\partial v}{\partial x}\right) + rS\frac{1}{S}K\frac{\partial v}{\partial x} - rKv = 0 \tag{6.21}$$

Eliminando S e S^2 além de K, temos:

$$-\frac{1}{2}\sigma^2\frac{\partial v}{\partial \zeta} + \frac{1}{2}\sigma^2\frac{\partial^2 v}{\partial x^2} - \frac{1}{2}\sigma^2\frac{\partial v}{\partial x} + r\frac{\partial v}{\partial x} - rv = 0 \tag{6.22}$$

Dividindo os dois membros da equação anterior por $-\frac{1}{2}\sigma^2$, sabendo que $\sigma > 0$, temos:

$$\frac{\partial v}{\partial \zeta} - \frac{\partial^2 v}{\partial x^2} + \frac{\partial v}{\partial x} - \frac{r}{\frac{1}{2}\sigma^2}\frac{\partial v}{\partial x} + \frac{r}{\frac{1}{2}\sigma^2}v = 0 \tag{6.23}$$

Isolando $\frac{\partial v}{\partial \zeta}$:

$$\frac{\partial v}{\partial \zeta} = \frac{\partial^2 v}{\partial x^2} + \left(\frac{r}{\frac{1}{2}\sigma^2} - 1\right)\frac{\partial v}{\partial x} - \frac{r}{\frac{1}{2}\sigma^2}v \tag{6.24}$$

Fazendo $k = \frac{r}{\frac{1}{2}\sigma^2}$:

$$\frac{\partial v}{\partial \zeta} = \frac{\partial^2 v}{\partial x^2} + (k-1)\frac{\partial v}{\partial x} - kv \tag{6.25}$$

Com a troca de variáveis, as condições iniciais modificam-se para:

$$\left\{\begin{array}{c} S=0 \Longrightarrow K\exp x = 0 \Longrightarrow x \to -\infty \\ \lim_{x\to-\infty} v(x,\zeta) = 0 \\ S \to \infty \Longrightarrow x \to \infty \\ \lim_{S\to\infty} C(S,t) \to S \Longrightarrow \lim_{x\to\infty} v(x,t) \to \infty \\ C(S,T) = \max(S-K,0) \Longrightarrow v(x,0) = \frac{1}{K}\max(S-K,0) = \\ \frac{1}{K}\max(K\exp x - K, 0) = \max(\exp x - 1, 0) \end{array}\right. \tag{6.26}$$

A equação (6.25) apresenta-se mais parecida com a equação de difusão do calor, porém ainda existem termos adicionais em $\frac{\partial v}{\partial x}$ e v. Vamos, então, considerar uma nova transformação:

$$v = \exp(\alpha x + \beta\zeta)u(x,\zeta) \tag{6.27}$$

onde α e β serão estabelecidos de forma apropriada para obtermos a equação diferencial (6.2) para a qual sabemos a solução.

Assim, vamos calcular as seguintes derivadas parciais.

Em relação a S:

$$\frac{\partial v}{\partial \zeta} = \exp(\alpha x + \beta\zeta)\beta u + \exp(\alpha x + \beta\zeta)\frac{\partial u}{\partial \zeta} =$$

$$\exp(\alpha x + \beta\zeta)\Big(\beta u + \frac{\partial u}{\partial \zeta}\Big) \tag{6.28}$$

Em relação a x:

$$\frac{\partial v}{\partial x} = \exp(\alpha x + \beta\zeta)\alpha u + \exp(\alpha x + \beta\zeta)\frac{\partial u}{\partial x} =$$

$$\exp(\alpha x + \beta\zeta)\Big(\alpha u + \frac{\partial u}{\partial x}\Big) \tag{6.29}$$

e

$$\frac{\partial^2 v}{\partial x^2} = \exp(\alpha x + \beta\zeta)\alpha\Big(\alpha u + \frac{\partial u}{\partial x}\Big) + \exp(\alpha x + \beta\zeta)\Big(\alpha\frac{\partial u}{\partial x} + \frac{\partial^2 u}{\partial x^2}\Big) =$$

$$\exp(\alpha x + \beta\zeta\Big)\Big(\alpha^2 u + 2\alpha\frac{\partial u}{\partial x} + \frac{\partial^2 u}{\partial x^2}\Big) \tag{6.30}$$

Substituindo (6.28), (6.29) e (6.30) em (6.25):

$$\exp(\alpha x + \beta\zeta)\Big(\beta u + \frac{\partial u}{\partial \zeta}\Big) =$$

$$\exp(\alpha x + \beta\zeta)\Big(\alpha^2 u + 2\alpha\frac{\partial u}{\partial x} + \frac{\partial^2 u}{\partial x^2}\Big) +$$

$$(k-1)\exp(\alpha x + \beta\zeta)\Big(\alpha u + \frac{\partial u}{\partial x}\Big) - k\exp(\alpha x + \beta\zeta)u \tag{6.31}$$

ou, eliminando $\exp(\alpha x + \beta\zeta)$:

$$\left(\beta u + \frac{\partial u}{\partial \zeta}\right) = \left(\alpha^2 u + 2\alpha\frac{\partial u}{\partial x} + \frac{\partial^2 u}{\partial x^2}\right) + (k-1)\left(\alpha u + \frac{\partial u}{\partial x}\right) - ku \tag{6.32}$$

Para obtermos uma equação da forma de (6.2) é necessário eliminar os termos em u e $\frac{\partial u}{\partial x}$. Para isto, basta escolher α e β apropriados. Assim, baseado na equação (6.32), devemos fazer:

$$\beta = \alpha^2 + (k-1)\alpha - k \tag{6.33}$$

e

$$2\alpha + (k-1) = 0 \tag{6.34}$$

Da equação (6.34), podemos extrair o valor de α. Assim:

$$\alpha = -\frac{(k-1)}{2} \tag{6.35}$$

Substituindo na equação (6.33) e isolando β, temos:

$$\beta = -\frac{(k+1)^2}{4} \tag{6.36}$$

Portanto, substituindo α e β na equação (6.27):

$$v = \exp\left(-\frac{1}{2}(k-1)x - \frac{1}{4}(k+1)^2\zeta\right)u(x,\zeta) \tag{6.37}$$

E com isso, garantimos que a equação (6.25) transforma-se em:

$$\frac{\partial u}{\partial \zeta} = \frac{\partial^2 u}{\partial x^2} \tag{6.38}$$

para $-\infty < x < \infty$ e $\zeta > 0$.

A condição inicial acompanha a mudança de variável. Assim, relembrando a equação (6.26), temos como condição inicial:

$$v(x,0) = \max(\exp(x) - 1, 0) \tag{6.39}$$

Substituindo $v(x,0)$ utilizando a condição de contorno para C em (6.26)

$$\exp(\alpha x + \beta 0)u(x,0) = \max(\exp(x) - 1, 0) \tag{6.40}$$

o que implica, isolando $u(x,0)$:

$$u(x,0) = \max\left[\exp\left(x + \frac{1}{2}(k-1)x\right) - \exp\left(\frac{1}{2}(k-1)x\right), 0\right] \tag{6.41}$$

Resolvendo,

$$u(x,0) = \max\left[\exp\left(\frac{1}{2}(k+1)x\right) - \exp\left(\frac{1}{2}(k-1)x\right), 0\right] \tag{6.42}$$

Assim, aplicando a solução da equação de difusão do calor (6.3), temos, considerando (6.38):

$$u(x,\zeta) = \frac{1}{2\sqrt{\pi\zeta}} \int_{-\infty}^{\infty} u_0(s) \exp\Big(-\frac{(x-s)^2}{4\zeta}\Big) ds \qquad (6.43)$$

onde
$u_0 = u(x,0)$ é dado pela equação (6.42).

Para calcular a integral acima, vamos realizar nova troca de variáveis:
Assim, fazendo:

$$y = \frac{(s-x)}{\sqrt{2\zeta}} \qquad (6.44)$$

obtemos

$$s = y\sqrt{2\zeta} + x \qquad (6.45)$$

e

$$dy = -\frac{1}{\sqrt{2\zeta}} ds \Longrightarrow ds = \sqrt{2\zeta} dy \qquad (6.46)$$

Além disso, os limites da integral transformam-se em:

$$\begin{cases} s \to \infty \Longrightarrow y \to \infty \\ s \to -\infty \Longrightarrow y \to -\infty \end{cases} \qquad (6.47)$$

Substituindo estes resultados na equação (6.43), temos:

$$u(x,\zeta) = \frac{1}{2\sqrt{\pi\zeta}} \int_{-\infty}^{\infty} u_0\Big(y\sqrt{2\zeta} + x\Big) \exp\Big(-\frac{(y^2\zeta)}{2\zeta}\Big) \sqrt{2\zeta} dy =$$
$$\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} u_0(y\sqrt{2\zeta} + x) \exp\Big(-\frac{1}{2}y^2\Big) dy \qquad (6.48)$$

Uma vez que não faz sentido $s \leq 0$, pois os preços de um ativo sempre são positivos, então:

$$s > 0 \Longrightarrow y > -\frac{x}{\sqrt{2\zeta}} \qquad (6.49)$$

e portanto, aplicando a equação referente a u_o (6.42), temos:

$$u(x,\zeta) = \frac{1}{\sqrt{2\pi}} \int_{-x/\sqrt{2\zeta}}^{\infty} \exp\Big[\frac{1}{2}(k+1)\Big(x + y\sqrt{2\zeta}\Big)\Big] \exp\Big(-\frac{1}{2}y^2\Big) dy -$$
$$\frac{1}{\sqrt{2\pi}} \int_{-x/\sqrt{2\zeta}}^{\infty} \exp\Big[\frac{1}{2}(k-1)\Big(x + y\sqrt{2\zeta}\Big)\Big] \exp\Big(-\frac{1}{2}y^2\Big) dy \qquad (6.50)$$

Assim,

$$u(x,\zeta) = I_1 - I_2 \qquad (6.51)$$

onde

$$I_1 = \frac{1}{\sqrt{2\pi}} \int_{-x/\sqrt{2\zeta}}^{\infty} \exp\Big[\frac{1}{2}(k+1)\Big(x + y\sqrt{2\zeta}\Big)\Big] \exp\Big(-\frac{1}{2}y^2\Big) dy \tag{6.52}$$

e

$$I_2 = \frac{1}{\sqrt{2\pi}} \int_{-x/\sqrt{2\zeta}}^{\infty} \exp\Big[\frac{1}{2}(k-1)\Big(x + y\sqrt{2\zeta}\Big)\Big] \exp\Big(-\frac{1}{2}y^2\Big) dy \tag{6.53}$$

Vamos portanto, calcular separadamente cada uma das parcelas acima.
Iniciemos pelo cálculo de I_1:

$$\begin{aligned} I_1 &= \frac{1}{\sqrt{2\pi}} \int_{-x/\sqrt{2\zeta}}^{\infty} \exp\Big(\frac{1}{2}(k+1)x - \frac{1}{2}(k+1)y\sqrt{2\zeta} - \frac{1}{2}y^2\Big) dy = \\ &= \frac{\exp\Big(\frac{1}{2}(k+1)x\Big)}{\sqrt{2\pi}} \int_{-x/\sqrt{2\zeta}}^{\infty} \exp\Big(-\frac{1}{2}(k+1)y\sqrt{2\zeta} - \frac{1}{2}y^2\Big) dy \end{aligned} \tag{6.54}$$

Para obtermos a integral acima, observemos que:

$$\begin{aligned} -\frac{1}{2}\Big(y - \frac{1}{2}(k+1)\sqrt{2\zeta}\Big)^2 &= -\frac{1}{2}\Big(y^2 - y(k+1)\sqrt{2\zeta} + \frac{1}{2}(k+1)^2\zeta\Big) \\ &= -\frac{1}{2}y^2 + \frac{1}{2}y(k+1)\sqrt{2\zeta} - \frac{1}{4}(k+1)^2\zeta \end{aligned} \tag{6.55}$$

Este fato nos sugere que completemos o quadrado na equação abaixo, multiplicando e dividindo a equação (6.55) por $\exp\Big(\frac{1}{4}(k+1)^2\zeta\Big)$.
Assim,

$$I_1 = \frac{\exp\Big(\frac{1}{2}(k+1)x + \frac{1}{4}(k+1)^2\zeta\Big)}{\sqrt{2\pi}} \int_{-x/\sqrt{2\zeta}}^{\infty} \exp\Big[-\frac{1}{2}\Big(y - \frac{1}{2}(k+1)\sqrt{2\zeta}\Big)^2\Big] dy \tag{6.56}$$

Fazendo $\rho = y - \frac{1}{2}(k+1)\sqrt{2\zeta}$, temos:

$$\left\{ \begin{array}{c} d\rho = dy \\ y \to \infty \Longrightarrow \rho \to \infty \\ y = -\frac{x}{\sqrt{2\zeta}} \Longrightarrow \rho = -\frac{x}{\sqrt{2\zeta}} - \frac{1}{2}(k+1)\sqrt{2\zeta} \end{array} \right. \tag{6.57}$$

Então, incorporando ρ na equação (6.56), temos:

$$I_1 = \frac{\exp\Big(\frac{1}{2}(k+1)x + \frac{1}{4}(k+1)^2\zeta\Big)}{\sqrt{2\pi}} \int_{-x/\sqrt{2\zeta}-(k+1)/2\sqrt{2\zeta}}^{\infty} \exp\Big(-\frac{1}{2}\rho^2\Big) d\rho \tag{6.58}$$

Passemos agora a investigar I_2:

$$I_2 = \frac{1}{\sqrt{2\pi}} \int_{-x/\sqrt{2\zeta}}^{\infty} \exp\Big[\frac{1}{2}(k-1)\Big(x + y\sqrt{2\zeta}\Big)\Big] \exp\Big(-\frac{1}{2}y^2\Big) dy$$

$$= \frac{\exp\left(\frac{1}{2}(k-1)x\right)}{\sqrt{2\pi}} \int_{-x/\sqrt{2\zeta}}^{\infty} \exp\left(-\frac{1}{2}(k-1)y\sqrt{2\zeta} - \frac{1}{2}y^2\right) dy \tag{6.59}$$

Novamente, observe que:

$$\begin{aligned} -\frac{1}{2}\left(y - \frac{1}{2}(k-1)\sqrt{2\zeta}\right)^2 &= -\frac{1}{2}\left(y^2 - y(k-1)\sqrt{2\zeta} + \frac{1}{2}(k-1)^2\zeta\right) \\ &= -\frac{1}{2}y^2 + \frac{1}{2}y(k+1)\sqrt{2\zeta} - \frac{1}{4}(k+1)^2\zeta \end{aligned} \tag{6.60}$$

Assim, para completar o quadrado, vamos multiplicar a equação (6.60) por $\exp\left(\frac{1}{4}(k-1)^2\zeta\right)$. Temos, então:

$$I_2 = \frac{\exp\left(\frac{1}{2}(k-1)x + \frac{1}{4}(k-1)^2\zeta\right)}{\sqrt{2\pi}} \int_{-x/\sqrt{2\zeta}}^{\infty} \exp\left[-\frac{1}{2}\left(y - \frac{1}{2}(k-1)\sqrt{2\zeta}\right)^2\right] dy \tag{6.61}$$

Fazendo $\rho = y - \frac{1}{2}(k-1)\sqrt{2\zeta}$, temos:

$$\begin{cases} d\rho = dy \\ y \to \infty \Longrightarrow \rho \to \infty \\ y = -\frac{x}{\sqrt{2\zeta}} \Longrightarrow \rho = -\frac{x}{\sqrt{2\zeta}} - \frac{1}{2}(k-1)\sqrt{2\zeta} \end{cases} \tag{6.62}$$

Então, incorporando ρ na equação (6.61), temos:

$$I_2 = \frac{\exp\left(\frac{1}{2}(k-1)x + \frac{1}{4}(k-1)^2\zeta\right)}{\sqrt{2\pi}} \int_{-x/\sqrt{2\zeta}-(k+1)/2\sqrt{2\zeta}}^{\infty} \exp\left(-\frac{1}{2}\rho^2\right) d\rho \tag{6.63}$$

Relembrando a fórmula da função de distribuição acumulada da normal padronizada

$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} \exp\left(-\frac{1}{2}s^2\right) ds \tag{6.64}$$

e notando que, pela simetria da distribuição normal,

$$\Phi(x) = 1 - \Phi(-x) \tag{6.65}$$

então podemos simplificar a notação de I_1 e I_2.

$$I_1 = \exp\left[\frac{1}{2}(k+1)x + \frac{1}{4}(k+1)^2\zeta\right]\Phi(d_1) \tag{6.66}$$

e

$$I_2 = \exp\left[\frac{1}{2}(k-1)x + \frac{1}{4}(k-1)^2\zeta\right]\Phi(d_2) \tag{6.67}$$

onde:

$$d_1 = \frac{x}{\sqrt{2\zeta}} + \frac{1}{2}(k+1)\sqrt{2\zeta} \tag{6.68}$$

$$\Phi(d_1) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{d_1} \exp\left(-\frac{1}{2}s^2\right) ds \tag{6.69}$$

e

$$d_2 = \frac{x}{\sqrt{2\zeta}} + \frac{1}{2}(k-1)\sqrt{2\zeta} \tag{6.70}$$

$$\Phi(d_2) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{d_2} \exp\left(-\frac{1}{2}s^2\right) ds \tag{6.71}$$

Como, por (6.51), $u(x,\zeta) = I_1 - I_2$, então podemos calcular $u(x,\zeta)$, substituindo os resultados obtidos em (6.66) e (6.67):

$$u(x,\zeta) = \exp\left(\frac{1}{2}(k+1)x + \frac{1}{4}(k+1)^2\zeta\right)\Phi(d_1) - \exp\left(\frac{1}{2}(k-1)x + \frac{1}{4}(k-1)^2\zeta\right)\Phi(d_2) \tag{6.72}$$

Desfazendo a troca de variáveis, podemos reconstruir $v(x,\zeta)$ a partir de $u(x,\zeta)$ considerando (6.27). Assim,

$$\begin{aligned} v(x,\zeta) = {} & \exp\left(-\frac{1}{2}(k-1)x - \frac{1}{4}(k+1)^2\zeta\right)\exp\left(\frac{1}{2}(k+1)x + \frac{1}{4}(k+1)^2\zeta\right)\Phi(d_1) \\ & - \exp\left(-\frac{1}{2}(k-1)x - \frac{1}{4}(k+1)^2\zeta\right)\exp\left(\frac{1}{2}(k-1)x + \frac{1}{4}(k-1)^2\zeta\right)\Phi(d_1) \end{aligned} \tag{6.73}$$

Colocando em evidência os termos comuns e reagrupando, temos:

$$\begin{aligned} v(x,\zeta) = {} & \exp\left[\left(-\frac{1}{2}k + \frac{1}{2} + \frac{1}{2}k + \frac{1}{2}\right)x\right]\Phi(d_1) \\ & - \exp\left[\left(-\frac{1}{4}k^2 - \frac{1}{2}k - \frac{1}{4} + \frac{1}{4}k^2 - \frac{1}{2}k + \frac{1}{4}\right)\zeta\right]\Phi(d_2) \end{aligned} \tag{6.74}$$

Podemos reduzir a equação acima a:

$$v(x,\zeta) = \exp(x)\Phi(d_1) - \exp(-k\zeta)\Phi(d_2) \tag{6.75}$$

Para acharmos o preço da opção, ou seja, $C(S,t)$, basta lembrarmos que dadas as condições de contorno (6.39), temos

$$\begin{cases} x = ln\left[\frac{S}{K}\right] \\ \zeta = \frac{1}{2}(T-t)\sigma^2 \\ v(x,\zeta) = \frac{C(S,t)}{K} \end{cases} \tag{6.76}$$

Então, substituindo estas equações em (6.75):

$$\frac{C(S,t)}{K} = \exp\left[\ln\left(\frac{S}{K}\right)\right]\Phi(d_1) - \exp\left[\left(-\frac{r}{\frac{1}{2}\sigma^2}\right)\frac{1}{2}(T-t)\sigma^2\right]\Phi(d_2) \tag{6.77}$$

Desenvolvendo a equação anterior, lembrando que ($\exp\ln x = x$) e tirando o mínimo múltiplo comum:

$$C(S,t) = S\Phi(d_1) - K\exp\Big[-r(T-t)\Big]\Phi(d_2) \tag{6.78}$$

onde d_1 e d_2 são dados por (6.68) e (6.70). Reescrevendo d_1 e d_2 como função das variáveis iniciais S e t, temos:

$$d_1 = \frac{\ln\frac{S}{K}}{\sqrt{2\frac{1}{2}(T-t)\sigma^2}} + \frac{1}{2}\left(\frac{r}{\frac{1}{2}\sigma^2}+1\right)\sqrt{2\frac{1}{2}(T-t)\sigma^2}$$

$$= \frac{\ln\left(\frac{S}{K}\right) + \frac{1}{2}\left(\frac{r}{\frac{1}{2}\sigma^2}+1\right)\sigma\sqrt{T-t}}{\sigma\sqrt{T-t}}$$

$$d_1 = \frac{\ln\left(\frac{S}{K}\right) + \frac{1}{2}\left(r+\frac{1}{2}\sigma^2\right)(T-t)}{\sigma\sqrt{T-t}} \tag{6.79}$$

Analogamente, obtemos d_2:

$$d_2 = \frac{\ln\frac{S}{K}}{\sqrt{2\frac{1}{2}(T-t)\sigma^2}} + \frac{1}{2}\left(\frac{r}{\frac{1}{2}\sigma^2}-1\right)\sqrt{2\frac{1}{2}(T-t)\sigma^2}$$

$$= \frac{\ln\left(\frac{S}{K}\right) + \frac{1}{2}\left(\frac{r}{\frac{1}{2}\sigma^2}-1\right)\sigma\sqrt{T-t}}{\sigma\sqrt{T-t}}$$

$$d_2 = \frac{\ln\left(\frac{S}{K}\right) + \frac{1}{2}\left(r-\frac{1}{2}\sigma^2\right)(T-t)}{\sigma\sqrt{T-t}} \tag{6.80}$$

As equações (6.78), (6.79) e (6.80) representam a fórmula analítica de Black e Scholes para precificação de opções de compra. Resumindo, temos que o valor da opção de compra no instante t, com vencimento em T, $T > t$ e preço de exercício K é dado por:

$$C(S,t) = S\Phi(d_1) - K\exp\Big[-r(T-t)\Big]\Phi(d_2) \tag{6.81}$$

com

$$d_1 = \frac{\ln\left(\frac{S}{K}\right) + \frac{1}{2}\left(r+\frac{1}{2}\sigma^2\right)(T-t)}{\sigma\sqrt{T-t}} \tag{6.82}$$

e

$$d_2 = \frac{\ln\left(\frac{S}{K}\right) + \frac{1}{2}\left(r-\frac{1}{2}\sigma^2\right)(T-t)}{\sigma\sqrt{T-t}} \tag{6.83}$$

6.3 Avaliação utilizando o conceito de neutralidade a risco

Utilizando a avaliação em uma situação em que os participantes do mercado são neutros ao risco, a fórmula de derivação de Black e Scholes torna-se um exercício de cálculo integral, simplificando sobremaneira a análise realizada na seção anterior. Primeiramente, apresentaremos o seguinte resultado que será utilizado e cuja demonstração encontra-se no Apêndice. Seja Y uma variável aleatória normalmente distribuída com média μ e variância σ^2, isto é:

$$Y \sim \mathcal{N}(\mu; \sigma^2)$$

Então, para todo $k \geq 0$,

$$E\Big[\exp(kY)1_{Y \geq a}\Big] = \int_a^{\infty} \exp(ky) dF_Y(y) = \exp\Big(k\mu + \frac{1}{2}k^2\sigma^2\Big)\Phi\Big(\frac{\mu - a + k\sigma^2}{\sigma}\Big) \tag{6.84}$$

onde

$F_Y(\cdot)$ é a função de distribuição da variável aleatória Y

$\Phi(\cdot)$ é a função normal padrão acumulada, isto é:

$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} \exp\Big(-\frac{z^2}{2} dz\Big) \tag{6.85}$$

com $x \in I\!R$.

Observemos também que $E\Big[\exp(kY)1_{Y \geq 0}\Big]$ pode ser interpretado como o k-ésimo momento truncado da variável aleatória $\exp(Y)$.

Vamos designar por S_t o preço da ação no instante t, com $t \in [0, T]$. Recordando o processo de difusão do preço da ação, para $t \in [0, T]$, temos:

$$S_t = S_0 \exp(\mu t + \sigma W_t) \tag{6.86}$$

Das equações (6.86) e (6.84), segue-se que o retorno da ação no tempo t, $\ln \frac{S_t}{S_0}$, tem distribuição normal com média $\Big(\mu + \frac{1}{2}\sigma^2\Big)t$ e variância $\sigma^2 t$.

Conforme discussão no Capítulo 2, no mundo risco-neutro todos os ativos têm o mesmo retorno esperado, que é igual à taxa de juros livre de risco r. Portanto, assumindo que os participantes do mercado são neutros ao risco, temos que:

$$\mu + \frac{1}{2}\sigma^2 = r$$

ou, equivalentemente,

$$\mu = r - \frac{1}{2}\sigma^2 \tag{6.87}$$

Agora, vamos calcular o valor esperado da opção no vencimento $t = T$:

$$E(C_T) = \int_K^{\infty} (x - K) dF_{S_T}(x) = \int_a^{\infty} (S_0 \exp(y) - K) dF_Y(y) \tag{6.88}$$

onde

$F_{S_T(\cdot)}$ é a função de distribuição de probabilidade de S_T
$F_Y(\cdot)$ é a função de distribuição de probabilidade de Y, com

$$Y \sim \mathcal{N}\Big(\Big((r - \frac{1}{2}\sigma^2\Big)T, \sigma^2 T\Big)$$

De acordo com (6.86) e (6.87), e considerando que a opção só será exercida se no vencimento $S_T > K$ então podemos fazer:

$$a = \ln \frac{K}{S_0} \tag{6.89}$$

Temos então:

$$E(C_T) = \int_{\ln \frac{K}{S_0}}^{\infty} S_0 \exp(y) dF_Y(y) - \int_{\ln \frac{K}{S_0}}^{\infty} K dF_Y(y) \tag{6.90}$$

Para utilizarmos o resultado (6.84), basta fixar para a primeira parcela acima $k = 1$ e para a segunda, $k = 0$. Assim,

$$\begin{aligned} E(C_T) = {} & S_0 \exp\Big[1\Big(r - \frac{1}{2}\sigma^2\Big)T - \frac{1}{2}1^2\sigma^2 T\Big]\Phi\Big[\frac{\Big(r - \frac{1}{2}\sigma^2\Big)T - \ln\frac{K}{S_0} + 1\sigma^2 T}{\sigma\sqrt{T}}\Big] \\ & - K \exp\Big[0\Big(r - \frac{1}{2}\sigma^2\Big)T - \frac{1}{2}0^2\sigma^2 T\Big]\Phi\Big[\frac{\Big(r - \frac{1}{2}\sigma^2\Big)T - \ln\frac{K}{S_0} + 0\sigma^2 T}{\sigma\sqrt{T}}\Big] \end{aligned} \tag{6.91}$$

Simplificando:

$$E(C_T) = S_0 \exp(rT)\Phi\Big[\frac{\ln\frac{S_0}{K} + \Big(r + \frac{1}{2}\sigma^2\Big)T}{\sigma\sqrt{T}}\Big] - K\Phi\Big[\frac{\ln\frac{S_0}{K} + \Big(r - \frac{1}{2}\sigma^2\Big)T}{\sigma\sqrt{T}}\Big] \tag{6.92}$$

O preço da opção no momento $t = 0$, C_0, é o valor esperado do prêmio no vencimento, $E(C_T)$, descontado pela taxa de juros livre de risco. Novamente, estamos utilizando o fato de que no mundo risco-neutro a taxa de retorno ou taxa de desconto de qualquer ativo corresponde a taxa de juros livre de risco.

Assim,

$$C_0 = \frac{E(C_T)}{\exp(rT)} \tag{6.93}$$

ou

$$C_0 = S_0\Phi(d_1) - K\exp(-rT)\Phi(d_2) \tag{6.94}$$

com

$$d_1 = \frac{\ln\frac{S_0}{K} + \Big(r + \frac{1}{2}\sigma^2\Big)T}{\sigma\sqrt{T}} \tag{6.95}$$

e

$$d_2 = \frac{\ln \frac{S_0}{K} + \left(r - \frac{1}{2}\sigma^2\right)T}{\sigma\sqrt{T}} \tag{6.96}$$

A equação (6.94) é equivalente à obtida na seção anterior (6.78) e portanto, podemos afirmar que a fórmula de Black e Scholes pode ser derivada através de duas abordagens bem diferentes. A primeira abordagem fundamenta-se na construção de uma carteira, com quantidades de ações e opções apropriadas, livre de risco e na suposição de que para não haver possibilidade de arbitragem, o retorno esperado da carteira deve ser igual à taxa de juros livre de risco. Nesta abordagem, são necessárias a utilização de conceitos de integral estocástica e a resolução de uma equação diferencial determinística.

A segunda abordagem baseia-se na suposição de que não há oportunidades de arbitragem e todos os participantes do mercado são neutros ao risco exigindo, independentemente do nível de risco assumido, sempre a taxa de juros livre de risco como retorno. Nesta abordagem, utilizando-se propriedade do k-ésimo momento da distribuição normal é necessário apenas o cálculo de uma integral.

Estes resultados, embora provenientes de origens diferentes, são compatíveis, tendo em vista que no processo de derivação da fórmula de Black e Scholes através da carteira livre de risco, a componente de retorno esperado do ativo μ desaparece devido à forma de construção da carteira e ao lema de Itô. Assim, o prêmio da opção independe da taxa de retorno esperada do ativo.

Com isso, a preferência pelo risco não é relevante e então qualquer estrutura de risco pode ser utilizada para precificar uma opção. Em particular, podemos escolher uma configuração onde os agentes do mercado sejam neutros ao risco. Esta assunção de neutralidade ao risco é conveniente tendo em vista a simplificação da derivação da fórmula de Black e Scholes. Cabe ressaltar que em diversos modelos de precificação o artifício de escolha de um mundo risco-neutro, facilita a obtenção de fórmulas analíticas.

Chapter 7

Método de busca de parâmetros de processos de mistura

A principal meta deste capítulo é a apresentação de um método estatístico de inferência dos parâmetros do processo que chamaremos de mistura simples. O método de mistura simples representa um caso particular do algoritmo EM (Expectation-Maximization) que está sendo amplamente utilizado pelos estatísticos. É bem conhecido o fato de que o algoritmo EM pode ser empregado para inferir os parâmetros de mistura de densidades, dado que as estruturas destas densidades sejam conhecidas. É fácil verificar que dados amostrais do mercado financeiro podem ser coletados de forma que o nosso problema de inferência nos parâmetros do processo de mistura simples seja reduzido a um problema de inferência dos parâmetros de uma mistura de densidades normais. Este problema foi estudado no artigo de Peters e Walker que fundamentará o estudo que se seguirá neste capítulo.

Como serão esclarecidas através da discussão apresentada na Seção 7.1, certas suposições são necessárias para que o método de Peters e Walker possa ser aplicado ao processo de mistura simples para que este modele adequadamente o comportamento de mercado. Uma vez que estas suposições podem não ser aceitáveis para algumas pessoas, poderia-se estudar a inferência dos parâmetros através do processo de mistura por processo de Poisson, tendo em vista sua melhor adequação à realidade.

Porém, não temos conhecimento sobre a existência de um método estatístico para tal tipo de inferência e o máximo que podemos fazer é mostrar as razões da não-aplicabilidade do método de Peters e Walker para a inferência de parâmetros de mistura por processos de Poisson.

Neste capítulo, vamos apresentar a teoria referente à estimação dos parâmetros do processo de mistura de brownianos. Inicialmente, na primeira seção, discutiremos as suposições que permitem aplicar o procedimento de Peters e Walker para a busca de parâmetros do processo de mistura de normais. Na segunda seção, abordaremos o estudo de Peters e Walker, calculando as passagens essenciais para a identificação do procedimento iterativo para a estimação através da maximização da função de verossimilhança. Na seção seguinte, vamos obter de forma análoga um outro procedimento iterativo que permite a imposição de uma restrição adicional para as médias das distribuições da mistura. Na quarta seção, obteremos as fórmulas recursivas para a obtenção dos parâmetros da mistura considerando-se que todos os processos componentes possuem a mesma variância. A principal contribuição deste capítulo é mostrar como a fórmula de Black e Scholes modificada, obtida no Capítulo 6, poderia ser utilizada na prática para a precificação de opções. Além disso, o material deste capítulo é apresentado de forma didática constituindo-se em uma fonte de referência para obtenção de uma fórmula de recorrência em situações de busca de parâmetros de uma mistura de brownianos sob diferentes restrições ou condições.

7.1 Suposições para aplicabilidade do procedimento de Peters e Walker

Vamos relembrar o processo S_t, $t \in [0,T]$, definido na Seção 4.1.2 e alternativamente, na Seção 4.1.3. Suponhamos que os valores dos parâmetros $\alpha_i, ..., \alpha_m, \mu_1, ..., \mu_m, \Sigma$ sejam desconhecidos. Observe que estamos, sem preda de generalidade, considerando processos com variâncias iguais. O objetivo desta seção é sugerir uma maneira de inferência destes parâmetros, baseados nas observações do processo.

Vamos partir da seguinte relação:

Se S_t, $t \in [0,T]$, é definido pela equação (6.12) então $X := \ln\left[\frac{S_1}{S_0}\right]$ tem a função de densidade $p(x)$, $x \in I\!R$, que satisfaz

$$p(x) = \sum_{i=1}^{m} \alpha_i p_i(x) \tag{7.1}$$

com $x \in I\!R$ e onde

$$p_i(x) = \frac{1}{\sqrt{2\pi\Sigma}} \exp\left(-\frac{1}{2}\frac{(x-\mu_i)^2}{\Sigma^2}\right) \tag{7.2}$$

sendo $x \in I\!R$ e $i = 1, \ldots, m$.

Nas próximas seções apresentaremos um método de busca de parâmetros da densidade $p(\cdot)$, que resolverá o nosso objetivo de modelamento através de um processo de mistura de brownianos. Sendo que este método baseia-se nas observações da variável aleatória X, a nossa meta é estabelecer uma maneira para a obtenção destas observações, partindo-se de observações de $S.$.

Para isto, vamos observar que é errado estabelecermos que

$$\ln\frac{S_1}{S_0}, \ln\frac{S_2}{S_1}, \cdots, \ln\frac{S_T}{S_{T-1}} \tag{7.3}$$

é uma sequência de variáveis aleatórias distribuídas como X e independentes. Este fato ocorre porque nossa suposição é , que uma vez que $S.$ escolheu um cenário, S_t o segue até o vencimento. Portanto, as variáveis aleatórias em (7.3) não são independentes.

O problema na retirada da amostra de X, como descrito acima, pode ser facilmente eliminado através de uma pequena mudança na interpretação de $S.$. Vamos supor que as notícias ou novas informações que alteram o comportamento da dinâmica do preço da ação chegam somente à noite, quando não há a realização de negociações com o ativo-objeto, por exemplo, ação de uma empresa, Vamos considerar que, de acordo com a ponderação de todas as notícias que chegaram durante a noite, logo na abertura do pregão na Bolsa de Valores no dia seguinte seja escolhido um determinado μ dentre os possíveis $\mu_1, \cdots, \mu_m$, que regerá o comportamento da ação durante todo o dia até o fechamento da Bolsa. Suporemos também que as notícias chegam de tal forma que implicará em um comportamento da ação no dia seguinte regido por $\mu_1, \cdots, \mu_m$ com as probabilidades de ocorrência $\alpha_1, \cdots, \alpha_m$, sendo que a escolha de determinado μ em um dia qualquer não depende da escolha em dias anteriores. Desta forma, o procedimento de Peters e Walker é adequado para a estimação dos parâmetros da mistura baseados em cotações dos preços das ações negociadas no mercado. Com esta interpretação, poderemos através dos dados amostrais a serem coletados no mercado finacneiro, inferir os parâmetros da mistura, que suporemos reger o comportamneto do preço do ativo-objeto.

Adicionalmente, serão discutidos neste trabalho modelos mais realistas do que o apresentado acima. Mostraremos que o método de busca de parâmetros que se aplica ao modelo descrito anteriormente, não pode ser utilizado para os outros modelos mais realistas que serão apresentados.

7.2 Procedimento de Peters e Walker

Seguindo o estudo de Peters e Walker, vamos estabelecer um método de inferência de parâmetros de uma função de densidade que é uma combinação convexa de densidades normais.

Seja m um número natural e sejam $\alpha_1, ..., \alpha_m$ m números reais que satisfazem:

$$\alpha_i^0 > 0 \tag{7.4}$$

com $i = 1, 2, ..., m$

e

$$\sum_{i=1}^{m} \alpha_i^0 = 1 \tag{7.5}$$

Sejam $(\mu_1^0, \Sigma_1^0), ..., (\mu_m^0, \Sigma_m^0)$ m pares de números reais tais que:

$$\Sigma_i^0 > 0 \tag{7.6}$$

$i = 1, 2, ..., m$

Designaremos por $p_i(\cdot)$ a função de densidade normal com média μ_i^0 e variância Σ_i^0:

$$p_i(x) = \frac{1}{\sqrt{2\pi\Sigma_i^0}} \exp\left(-\frac{1}{2}\frac{(x-\mu_i)^2}{\Sigma_i^0}\right) \tag{7.7}$$

$x \in I\!R$, $i = 1, ..., m$

Definiremos a função $p(\cdot)$ como

$$p(x) = \sum_{i=1}^{m} \alpha_i^0 p_i(x) \tag{7.8}$$

com $x \in I\!R$.

Pela definição acima, $p(x)$ é uma função de densidade de probabilidades. Vamos denotar por X a variável aleatória cuja função de densidade de probabilidades é $p(\cdot)$. O problema abordado neste capítulo refere-se à estimação dos valores de $\alpha_i^0, ..., \alpha_m^0, \mu_1^0, ..., \mu_m^0, \Sigma_1^0, ..., \Sigma_m^0$ com base nas observações de X. Estas observações serão representadas a seguir por $x_1, ..., x_N$.

Cabe observar que não incluímos m como parâmetro a ser estimado, seguindo o procedimento de Peters e Walker que não estabelece um critério para a determinação da quantidade de movimentos brownianos presentes no processo de mistura.

Discutiremos agora a relação deste procedimento de busca de parâmetros com o problema da estimativa dos parâmetros do processo que foi sugerido no Capítulo 4 para representar o modelo da dinâmica de formação do preço da ação.

Sejam agora $x_1, x_2, ..., x_N$ N observações independentes da variável aleatória X. Vamos definir a função de verossimilhança:

$$f(x_1, ..., x_N) = \prod_{k=1}^{N} p(x_k) \tag{7.9}$$

Observemos que a função definida acima é uma função de $\alpha_1^0, ..., \alpha_m^0, \mu_1^0, ..., \mu_m^0, \Sigma_1^0, ..., \Sigma_m^0$. Denotaremos por $\alpha_1, ...\alpha_m, \mu_1, ..., \mu_m, \Sigma_1, ..., \Sigma_m$ os valores que maximizam a função f. Estes valores são as estimativas dos verdadeiros parâmetros $\alpha_1^0, ..., \alpha_m^0, \mu_1^0, ..., \mu_m^0, \Sigma_1^0, ..., \Sigma_m^0$. Estas estimativas são chamadas de estimativas de máxima verossimilhança. São soluções de um sistema de equações que expressam a condição de maximização da função f. Porém, em nosso caso, este sistema de equações não possui solução fechada. Portanto, obteremos a solução através de um procedimento que gera, recursivamente, uma sequência de vetores que convergem à solução do sistema. Apresentaremos a seguir a derivação deste procedimento.

Vamos estabelecer uma variável aleatória x cuja função densidade de probabilidade é uma combinação convexa de densidades normais. Assim:

$$p(x) = \sum_{i=1}^{m} \alpha_i^0 p_i(x) \tag{7.10}$$

onde

$$x \in I\!R$$

$$\alpha_i^0 > 0 \tag{7.11}$$

$i = 1, 2, \cdots, m$ e

$$\sum_{i=1}^{m} \alpha_i^0 = 1 \tag{7.12}$$

$$p_i(x) = \frac{1}{\sqrt{2\pi}(\Sigma_i^0)^{1/2}} \exp\Big[-\frac{1}{2}(x - \mu_i^0)^2 (\Sigma_i^0)^{-1}\Big] \tag{7.13}$$

Observe que os α_i^0 correspondem aos pesos de cada distribuição normal na mistura. Em outras palavras, α_i^0 são as probabilidades de os retornos do ativo seguirem cada um dos brownianos que compõem o processo de mistura e portanto, as restrições denotadas pelas equações (7.11) e (7.12) são necessárias.

Seja $x_k \in I\!R$, com $k = 1, ..., N$, uma amostra independente de observações da variável aleatória X. A função de verossimilhança é dada por:

$$f(x_1, ..., x_N) = \prod_{k=1}^{N} p(x_k) \tag{7.14}$$

Com a finalidade de facilitar os cálculos, podemos aplicar o logaritmo nos membros da equação acima. Assim, teremos:

$$L(x_1, ..., x_N) = \ln[f(x_1, ..., x_N)] = \ln\Big[\prod_{k=1}^{N} p(x_k)\Big] = \sum_{k=1}^{N} \ln[p(x_k)] = \sum_{k=1}^{N} \ln\Big[\sum_{i=1}^{m} \alpha_i^0 p_i(x_k)\Big] \tag{7.15}$$

As estimativas de máxima verossimilhança dos verdadeiros parâmetros $(\alpha_i^0, \mu_i^0, \Sigma_i^0)_{i=1,...,m}$ são obtidas através das escolhas dos parâmetros $(\alpha_i, \mu_i, \Sigma_i)_{i=1,...,m}$ que maximizam localmente a função de verossimilhança (7.14) ou, equivalentemente, a função descrita por (7.15).

Note, portanto, que p deve ser avaliado trocando-se os verdadeiros parâmetros $(\alpha_i^0, \mu_i^0, \Sigma_i^0)_{i=1,...,m}$ pelas estimativas $(\alpha_i, \mu_i, \Sigma_i)_{i=1,...,m}$. Além disso, a busca de um máximo local evita as dificuldades que podem surgir devido à possibilidade de L não possuir máximo global.

Para achar o ponto de máximo, uma vez que a função L é diferenciável com relação aos parâmetros a serem estimados, podemos utilizar a função de Lagrange F onde à função de verossimilhança é adicionada uma parcela referente à restrição imposta por (7.12). Assim,

$$F(\alpha_1, ..., \alpha_m, \mu_1, ..., \mu_m, \Sigma_1, ..., \Sigma_m) = L + \lambda(\sum_{i=1}^{m} \alpha_i - 1) \tag{7.16}$$

Derivando parcialmente F em relação aos parâmetros $(\alpha_i, \mu_i, \Sigma_i)_{i=1,...,m}$ e em relação a λ podemos igualar cada derivada parcial a zero para obtemos um ponto de máximo. Assim, derivando em relação a λ:

$$\frac{\partial F}{\partial \lambda} = 0 \Longrightarrow \sum_{i=1}^{m} \alpha_i - 1 = 0 \Longrightarrow \sum_{i=1}^{m} \alpha_i = 1 \tag{7.17}$$

Derivando em relação a α_i, temos:

$$\frac{\partial F}{\partial \alpha_i} = 0 \Longrightarrow \sum_{k=1}^{N} \frac{1}{\sum_{i=1}^{m} \alpha_i p_i(x_k)} p_i(x_k) + \lambda = 0 \tag{7.18}$$

Aplicando (7.10) em (7.18):

$$\sum_{k=1}^{N} \frac{p_i(x_k)}{p(x_k)} = -\lambda \tag{7.19}$$

A equação acima nos permite escrever:

$$\begin{cases} \alpha_1 \sum_{k=1}^{N} \frac{p_1(x_k)}{p(x_k)} = -\alpha_1 \lambda \\ \vdots \\ \alpha_m \sum_{k=1}^{N} \frac{p_m(x_k)}{p(x_k)} = -\alpha_m \lambda \end{cases} \tag{7.20}$$

E portanto, somando as equações acima membro a membro, temos:

$$\alpha_1 \sum_{k=1}^{N} \frac{p_1(x_k)}{p(x_k)} + ... + \alpha_m \sum_{k=1}^{N} \frac{p_m(x_k)}{p(x_k)} = -\alpha_1 \lambda - ... - \alpha_m \lambda \tag{7.21}$$

Ou,

$$\sum_{k=1}^{N} \alpha_1 \frac{p_1(x_k)}{p(x_k)} + ... + \sum_{k=1}^{N} \alpha_m \frac{p_m(x_k)}{p(x_k)} = -(\alpha_1 + ... + \alpha_m)\lambda \tag{7.22}$$

Aplicando (7.12), isto é, $\sum_{\alpha_i} = 1$ na equação acima, temos:

$$\sum_{k=1}^{N} \frac{\alpha_1 p_1(x_k) + ... + \alpha_m p_m(x_k)}{p(x_k)} = -\lambda \tag{7.23}$$

Utilizando (7.10), sabemos que $p(x) = \sum_{i=1}^{m} \alpha_i^0 p_i(x)$, então:

$$\sum_{k=1}^{N} \frac{p(x_k)}{p(x_k)} = -\lambda \tag{7.24}$$

E portanto:

$$N = -\lambda \tag{7.25}$$

Para encontramos uma fórmula de recorrência para α_i, vamos utilizar um artifício matemático multiplicando e dividindo a equação (7.19) por α_i e substituindo λ por $-N$. Assim, isolando α_i, temos:

$$\alpha_i = \frac{\alpha_i}{N} \sum_{k=1}^{N} \frac{p_i(x_k)}{p(x_k)} \tag{7.26}$$

Vamos agora obter uma fórmula de recorrência para μ_i, derivando F em relação a μ_i.

$$\frac{\partial F}{\partial \mu_i} = 0 \Longrightarrow \sum_{k=1}^{N} \frac{p_i(x_k)}{\sum_{i=1}^{m} \alpha_i p_i(x_k)} \left[-\frac{1}{2} 2(x_k - \mu_i)(-1)\Sigma_i^{(-1)} \right] = 0 \tag{7.27}$$

Por (7.13) e sabendo que $\Sigma_i > 0$:

$$\sum_{k=1}^{N} \frac{p_i(x_k)}{p(x_k)} \frac{(x_k - \mu_i)}{\Sigma_i} = 0 \Longrightarrow \sum_{k=1}^{N} x_k \frac{p_i(x_k)}{p(x_k)} - \sum_{k=1}^{N} \mu_i \frac{p_i(x_k)}{p(x_k)} = 0 \tag{7.28}$$

Isolando μ_i:

$$\mu_i = \frac{\sum_{k=1}^{N} x_k \frac{p_i(x_k)}{p(x_k)}}{\sum_{k=1}^{N} \frac{p_i(x_k)}{p(x_k)}} \tag{7.29}$$

Para fins de acuidade computacional, multiplica-se e divide-se o segundo membro da equação acima por $1/N$. Assim, obtemos a seguinte forma de recorrência para μ_i:

$$\mu_i = \frac{\frac{1}{N} \sum_{k=1}^{N} x_k \frac{p_i(x_k)}{p(x_k)}}{\frac{1}{N} \sum_{k=1}^{N} \frac{p_i(x_k)}{p(x_k)}} \tag{7.30}$$

Vamos agora derivar F em relação a Σ_i

$$\frac{\partial F}{\partial \Sigma_i} = 0 \Longrightarrow$$

$$\sum_{k=1}^{N} \frac{1}{\sum_{i=1}^{m} \alpha_i p_i(x_k)} \left[\alpha_i \frac{1}{\sqrt{2\pi}} \left(-\frac{1}{2} \Sigma_i^{-3/2} \exp\left(-\frac{1}{2}(x_k - \mu_i)^2 \Sigma_i^{-1} \right) \right) + \right.$$

$$\left(\Sigma^{-1/2}\exp\left(-\frac{1}{2}(x_k-\mu_i)^2\Sigma_i^{-1}\right)\left(-\frac{1}{2}(x_k-\mu_i)^2(-1)\Sigma^{-2}\right)\right)\Bigg] = 0 \tag{7.31}$$

Ou seja,

$$\sum_{k=1}^{N}\frac{\alpha_i\frac{1}{\sqrt{2\pi}}\Sigma_i^{-1/2} exp(-\frac{1}{2}(x_k-\mu)^2\Sigma^{-1}}{p(x_k)}\left[-\frac{1}{2}\Sigma_i^{-1}+\frac{1}{2}\Sigma_i^{-2}(x_k-\mu_i)^2\right] = 0 \tag{7.32}$$

Por (7.13) e sabendo que $\Sigma_i > 0$:

$$\sum_{k=1}^{N}\frac{p_i(x_k)}{p(x_k)}[-\Sigma_i + (x_k-\mu_i)^2] = 0 \tag{7.33}$$

Assim, rearranjando os termos:

$$\sum_{k=1}^{N}\frac{p_i(x_k)}{p(x_k)}\Sigma_i = \sum_{k=1}^{N}\frac{p_i(x_k)}{p(x_k)}(x_k-\mu_i)^2 \tag{7.34}$$

Isolando Σ_i e multiplicando e dividindo por $1/N$:

$$\Sigma_i = \frac{\frac{1}{N}\sum_{k=1}^{N}\frac{p_i(x_k)}{p(x_k)}(x_k-\mu_i)^2}{\frac{1}{N}\sum_{k=1}^{N}\frac{p_i(x_k)}{p(x_k)}} \tag{7.35}$$

De acordo com os resultados (7.26), (7.30) e (7.35), podemos então resumir o procedimento iterativo para a obtenção das estimativas dos parâmetros da mistura $\alpha_i, ..., \alpha_m, \mu_1, ..., \mu_m, \Sigma_1, ..., \Sigma_m$. A partir de um conjunto de valores iniciais, também denominado chute inicial, $\alpha_i(0), ..., \alpha_m(0)$, $\mu_1(0), ..., \mu_m(0)$, $\Sigma_1(0), ..., \Sigma_m(0)$, são calculados os próximos conjuntos de parâmetros $\alpha_i(j), ..., \alpha_m(j), \mu_1(j), ..., \mu_m(j), \Sigma_1(j$ com $j \in I\!N^*$, utilizando-se as seguintes fórmulas de recorrência:

$$\alpha_i(j) = \frac{\alpha_i(j-1)}{N}\sum_{k=1}^{N}\frac{p_i(x_k)}{p(x_k)} \tag{7.36}$$

$$\mu_i(j) = \frac{\frac{1}{N}\sum_{k=1}^{N} x_k\frac{p_i(x_k)}{p(x_k)}}{\frac{1}{N}\sum_{k=1}^{N}\frac{p_i(x_k)}{p(x_k)}} \tag{7.37}$$

$$\Sigma_i(j) = \frac{\frac{1}{N}\sum_{k=1}^{N}\frac{p_i(x_k)}{p(x_k)}(x_k-\mu_i(j-1))^2}{\frac{1}{N}\sum_{k=1}^{N}\frac{p_i(x_k)}{p(x_k)}} \tag{7.38}$$

com $p_i(x_k)$ e $p(x_k)$ sendo funções de $\alpha_1(j-1), ..., \alpha_m(j-1), \mu_1(j-1), ...\mu_m(j-1), \Sigma_1(j-1), ..., \Sigma_m(j-1)$.

O procedimento é repetido tantas vezes quanto forem necessárias para a obediência a um critério de convergência predeterminado.

7.3 Procedimento de Peters e Walker modificado - restrição para médias

Vamos agora desenvolver as fórmulas de recorrência para obtenção dos parâmetros da mistura, incorporando uma nova restrição. Assim, vamos procurar os parâmetros da mistura de forma que a restrição abaixo, referente às médias, seja também obedecida:

$$\sum_{i=1}^{m} \alpha_i \mu_i = \gamma$$

onde γ representa uma tendência determinística dos retornos do ativo, que pode ser fixada. Temos assim, de modo análogo ao exposto na seção anterior:

$$p(x) = \sum_{i=1}^{m} \alpha_i^0 p_i(x) \tag{7.39}$$

onde

$$x \in \mathbb{R}$$

$$i = 1, ..., m$$

$$\alpha_i^0 > 0 \tag{7.40}$$

$$\sum_{i=1}^{m} \alpha_i^0 = 1 \tag{7.41}$$

$$\sum_{i=1}^{m} \alpha_i \mu_i = 0 \tag{7.42}$$

$$p_i(x) = \frac{1}{\sqrt{2\pi}(\Sigma_i^0)^{1/2}} \exp\left[-\frac{1}{2}(x - \mu_i^0)^2 (\Sigma_i^0)^{-1}\right] \tag{7.43}$$

Seja $x_k \in \mathbb{R}$, com $k = 1, ..., N$, uma amostra independente de observações da variável aleatória x. A função de Lagrange, já tendo sido aplicado o logaritmo do função de verossimilhança e incorporada a restrição (7.42), é dada por:

$$F(\alpha_1, ..., \alpha_m, \mu_1, ..., \mu_m, \Sigma_1, ..., \Sigma_m) = L + \lambda_1(\sum_{i=1}^{m} \alpha_i - 1) + \lambda_2(\sum_{i=1}^{m} \alpha_i \mu_i - \gamma) \tag{7.44}$$

Vamos derivar F em relação aos parâmetros $(\alpha_i, \mu_i, \Sigma_i)_{i=1,...,m}$ e em relação a λ_1 e a λ_2 para posteriormente igualar a zero as derivadas parciais para obter um ponto de máximo.

Assim, derivando F em relação a λ_1:

$$\frac{\partial F}{\partial \lambda_1} = 0 \Longrightarrow \sum_{i=1}^{m} \alpha_i - 1 = 0 \Longrightarrow \sum_{i=1}^{m} \alpha_i = 1 \tag{7.45}$$

Derivando F em relação a λ_2

$$\frac{\partial F}{\partial \lambda_2} = 0 \Longrightarrow \sum_{i=1}^{m} \alpha_i \mu_i = \gamma \Longrightarrow \sum_{i=1}^{m} \alpha_i \mu_i = \gamma \tag{7.46}$$

Derivando agora F em relação a α_i, temos:

$$\frac{\partial F}{\partial \alpha_i} = 0 \Longrightarrow \sum_{k=1}^{N} \frac{1}{\sum_{i=1}^{m} \alpha_i p_i(x_k)} p_i(x_k) + \lambda_1 + \lambda_2 \mu_i = 0 \tag{7.47}$$

Aplicando (7.39) em (7.47):

$$\sum_{k=1}^{N} \frac{p_i(x_k)}{p(x_k)} + \lambda_1 + \lambda_2 \mu_i = 0 \tag{7.48}$$

Isolando $-\lambda_1$, temos:

$$\sum_{k=1}^{N} \frac{p_i(x_k)}{p(x_k)} + \lambda_2 \mu_i = -\lambda_1 \tag{7.49}$$

E portanto, a equação acima pode ser escrita para cada i, multiplicando os dois membros por α_i:

$$\begin{cases} \alpha_1 \sum_{k=1}^{N} \frac{p_1(x_k)}{p(x_k)} + \alpha_1 \lambda_2 \mu_1 = -\alpha_1 \lambda_1 \\ \vdots \\ \alpha_m \sum_{k=1}^{N} \frac{p_m(x_k)}{p(x_k)} + \alpha_m \lambda_2 \mu_m = -\alpha_m \lambda_1 \end{cases} \tag{7.50}$$

Somando as equações anteriores membro a membro, obtemos:

$$\alpha_1 \sum_{k=1}^{N} \frac{p_1(x_k)}{p(x_k)} + \ldots + \alpha_m \sum_{k=1}^{N} \frac{p_m(x_k)}{p(x_k)} + \alpha_1 \lambda_2 \mu_1 + \ldots + \alpha_m \lambda_2 \mu_m = -\alpha_1 \lambda_1 - \ldots - \alpha_m \lambda_1 \tag{7.51}$$

Colocando λ_1 e λ_2 em evidência:

$$\sum_{k=1}^{N} \alpha_1 \frac{p_1(x_k)}{p(x_k)} + \ldots + \sum_{k=1}^{N} \alpha_m \frac{p_m(x_k)}{p(x_k)} + \lambda_2(\alpha_1 \mu_1 + \ldots + \alpha_m \mu_m) = -(\alpha_1 + \ldots + \alpha_m)\lambda_1 \tag{7.52}$$

Aplicando (7.41) e (7.42) na equação acima:

$$\sum_{k=1}^{N} \frac{\alpha_1 p_1(x_k) + \ldots + \alpha_m p_m(x_k)}{p(x_k)} = -\lambda_1 \tag{7.53}$$

Por (7.39):

$$\sum_{k=1}^{N} \frac{p(x_k)}{p(x_k)} = -\lambda_1 \tag{7.54}$$

E portanto:

$$N = -\lambda_1 \tag{7.55}$$

Derivando F em relação a μ_i, temos:

$$\frac{\partial F}{\partial \mu_i} = 0 \Longrightarrow \sum_{k=1}^{N} \frac{\alpha_i p_i(x_k)}{\sum_{i=1}^{m} \alpha_i p_i(x_k)} \left[-\frac{1}{2} 2(x_k - \mu_i)(-1)\Sigma_i^{-1} \right] + \lambda_2 \alpha_i = 0 \tag{7.56}$$

Sabendo que $\alpha_i > 0$

$$\sum_{k=1}^{N} \frac{p_i(x_k)}{\sum_{i=1}^{m} \alpha_i p_i(x_k)} (x_k - \mu_i) \Sigma_i^{-1} = -\lambda_2 \tag{7.57}$$

Aplicando novamente 7.39 na equação acima:

$$\sum_{k=1}^{N} \frac{p_i(x_k)}{p(x_k)} \frac{(x_k - \mu_i)}{\Sigma_i} = -\lambda_2 \tag{7.58}$$

A equação acima nos permite calcular para cada i, multiplicando os dois membros por α_i:

$$\left\{ \begin{array}{c} \alpha_1 \sum_{k=1}^{N} \frac{p_1(x_k)}{p(x_k)} \frac{(x_k - \mu_1)}{\Sigma_1} = -\alpha_1 \lambda_2 \\ \vdots \\ \alpha_m \sum_{k=1}^{N} \frac{p_m(x_k)}{p(x_k)} \frac{(x_k - \mu_m)}{\Sigma_m} = -\alpha_m \lambda_2 \end{array} \right. \tag{7.59}$$

Somando as equações acima membro a membro:

$$\alpha_1 \sum_{k=1}^{N} \frac{p_1(x_k)}{p(x_k)} \frac{(x_k - \mu_1)}{\Sigma_1} + \ldots + \alpha_m \sum_{k=1}^{N} \frac{p_m(x_k)}{p(x_k)} \frac{(x_k - \mu_m)}{\Sigma_m} = -\lambda_2(\alpha_1 + \ldots + \alpha_m) \tag{7.60}$$

Utilizando a equação (7.41) em (7.60):

$$\sum_{k=1}^{N} \frac{\alpha_1 p_1(x_k)}{p(x_k)} \frac{(x_k - \mu_1)}{\Sigma_1} + \ldots + \sum_{k=1}^{N} \frac{\alpha_m p_m(x_k)}{p(x_k)} \frac{(x_k - \mu_m)}{\Sigma_m} = -\lambda_2 \tag{7.61}$$

E assim,

$$\sum_{k=1}^{N} \sum_{i=1}^{m} \frac{\alpha_i p_i(x_k)}{p(x_k)} \frac{(x_k - \mu_i)}{\Sigma_i} = -\lambda_2 \tag{7.62}$$

Até este momento então, obtivemos os valores de λ_1 e λ_2 dados pelas equações (7.55) e (7.62). Observando a igualdade das equações (7.58) e (7.62), temos:

$$\sum_{k=1}^{N} \frac{p_i(x_k)}{p(x_k)} \frac{(x_k - \mu_i)}{\Sigma_i} = \sum_{k=1}^{N} \sum_{i=1}^{m} \frac{\alpha_i p_i(x_k)}{p(x_k)} \frac{(x_k - \mu_i)}{\Sigma_i} \tag{7.63}$$

Desenvolvendo a equação acima:

$$\sum_{k=1}^{N} \frac{p_i(x_k)}{p(x_k)} \frac{(x_k)}{\Sigma_i} - \sum_{k=1}^{N} \frac{p_i(x_k)}{p(x_k)} \frac{(\mu_i)}{\Sigma_i} = \sum_{k=1}^{N} \sum_{i=1}^{m} \frac{\alpha_i p_i(x_k)}{p(x_k)} \frac{(x_k - \mu_i)}{\Sigma_i} \tag{7.64}$$

Sabendo que $\Sigma_i > 0$

$$\mu_i \sum_{k=1}^{N} \frac{p_i(x_k)}{p(x_k)} = \sum_{k=1}^{N} \frac{x_k p_i(x_k)}{p(x_k)} - \Sigma_i \sum_{k=1}^{N} \sum_{i=1}^{m} \frac{\alpha_i p_i(x_k)}{p(x_k)} \frac{(x_k - \mu_i)}{\Sigma_i} \tag{7.65}$$

Isolando μ_i e multiplicando e dividindo por $1/N$, temos:

$$\mu_i = \frac{\frac{1}{N} \sum_{k=1}^{N} \frac{x_k p_i(x_k)}{p(x_k)} - \frac{1}{N} \Sigma_i \sum_{k=1}^{N} \sum_{i=1}^{m} \frac{\alpha_i p_i(x_k)}{p(x_k)} \frac{(x_k - \mu_i)}{\Sigma_i}}{\frac{1}{N} \sum_{k=1}^{N} \frac{p_i(x_k)}{p(x_k)}} \tag{7.66}$$

Substituindo (7.55), (7.62) e (7.66) em (7.49):

$$\sum_{k=1}^{N} \frac{p_i(x_k)}{p(x_k)} - \left(\sum_{k=1}^{N} \frac{p_i(x_k)}{p(x_k)} \frac{(x_k - \mu_i)}{\Sigma_i}\right)$$
$$\left(\frac{\frac{1}{N}\sum_{k=1}^{N} \frac{x_k p_i(x_k)}{p(x_k)} - \frac{1}{N}\Sigma_i \sum_{k=1}^{N}\sum_{i=1}^{m} \frac{\alpha_i p_i(x_k)}{p(x_k)}\frac{(x_k-\mu_i)}{\Sigma_i}}{\frac{1}{N}\sum_{k=1}^{N}\frac{p_i(x_k)}{p(x_k)}}\right) = N \tag{7.67}$$

Multiplicando os dois membros por α_i, podemos rearranjar a equação acima para encontrar uma fórmula de recorrência para α_i:

$$\alpha_i = \frac{\alpha_i}{N}\sum_{k=1}^{N}\frac{p_i(x_k)}{p(x_k)} -$$
$$\frac{\alpha_i}{N}\left(\sum_{k=1}^{N} \frac{p_i(x_k)}{p(x_k)} \frac{(x_k - \mu_i)}{\Sigma_i}\right)\left(\frac{\frac{1}{N}\sum_{k=1}^{N} \frac{x_k p_i(x_k)}{p(x_k)} - \frac{1}{N}\Sigma_i \sum_{k=1}^{N}\sum_{i=1}^{m} \frac{\alpha_i p_i(x_k)}{p(x_k)}\frac{(x_k-\mu_i)}{\Sigma_i}}{\frac{1}{N}\sum_{k=1}^{N}\frac{p_i(x_k)}{p(x_k)}}\right) \tag{7.68}$$

Derivando F em relação a Σ_i e igualando a zero, temos:

$$\frac{\partial F}{\partial \Sigma_i} = 0 \Longrightarrow$$

$$\sum_{k=1}^{N} \frac{1}{\sum_{i=1}^{m}\alpha_i p_i(x_k)}\left(\alpha_i \frac{1}{\sqrt{2\pi}}\right)\left[-\frac{1}{2}\Sigma^{-3/2}\exp\left(-\frac{1}{2}(x_k-\mu_i)^2\Sigma^{-1}\right) + \right.$$
$$\left.\Sigma^{-1/2}\exp\left(-\frac{1}{2}(x_k-\mu_i)^2\Sigma^{-1}\right).\left(-\frac{1}{2}(x_k-\mu_i)^2(-1)\Sigma^{-2}\right)\right] = 0 \tag{7.69}$$

Ou seja,

$$\sum_{k=1}^{N} \frac{\alpha_i \frac{1}{\sqrt{2\pi}}\Sigma^{-1/2}\exp(-\frac{1}{2}(x_k-\mu)^2\Sigma^{-1})}{p(x_k)}\left[-\frac{1}{2}\Sigma^{-1} + \frac{1}{2}\Sigma^{-2}(x_k-\mu_i)^2\right] = 0 \tag{7.70}$$

Por (7.43) e sabendo que $\Sigma_i > 0$:

$$\sum_{k=1}^{N}\frac{p_i(x_k)}{p(x_k)}[\Sigma_i + (x_k-\mu_i)^2] = 0 \tag{7.71}$$

Isto implica

$$\sum_{k=1} N\frac{p_i(x_k)}{p(x_k)}\Sigma_i = \sum_{k=1} N\frac{p_i(x_k)}{p(x_k)}(x_k-\mu_i)^2 \tag{7.72}$$

Isolando Σ_i e multiplicando e dividindo por $1/N$:

$$\Sigma_i = \frac{\frac{1}{N}\sum_{k=1}^{N}\frac{p_i(x_k)}{p(x_k)}(x_k-\mu_i)^2}{\frac{1}{N}\sum_{k=1}^{N}\frac{p_i(x_k)}{p(x_k)}} \tag{7.73}$$

Novamente, considerando as equações (7.68), (7.66) e (7.73), podemos utilizar o procedimento iterativo para a obtenção das estimativas dos parâmetros da mistura $\alpha_i, ..., \alpha_m, \mu_1, ..., \mu_m, \Sigma_1, ..., \Sigma_m$. Dado o chute inicial $\alpha_i(0), ..., \alpha_m(0), \mu_1(0), ..., \mu_m(0), \Sigma_1(0), ..., \Sigma_m(0)$, pode-se obter os próximos conjuntos de parâmetros $\alpha_i(j), ..., \alpha_m(j), \mu_1(j), ..., \mu_m(j), \Sigma_1(j), ..., \Sigma_m(j)$, com $j \in I\!N^*$, através das seguintes fórmulas de recorrência:

$$\alpha_i(j) = \frac{\alpha_i(j-1)}{N} \sum_{k=1}^{N} \frac{p_i(x_k)}{p(x_k)} -$$

$$\frac{\alpha_i(j-1)}{N}\Big(\sum_{k=1}^{N} \frac{p_i(x_k)}{p(x_k)} \frac{(x_k - \mu_i(j-1))}{\Sigma_i(j-1)}\Big)\Big(\frac{\frac{1}{N}\sum_{k=1}^{N} \frac{x_k p_i(x_k)}{p(x_k)} - \frac{1}{N}\Sigma_i(j-1)\sum_{k=1}^{N}\sum_{i=1}^{m} \frac{\alpha_i(j-1)p_i(x_k)}{p(x_k)} \frac{(x_k-\mu_i(j-1))}{\Sigma_i(j-1)}}{\frac{1}{N}\sum_{k=1}^{N} \frac{p_i(x_k)}{p(x_k)}}\Big) \tag{7.74}$$

$$\mu_i(j) = \frac{\frac{1}{N}\sum_{k=1}^{N} \frac{x_k p_i(x_k)}{p(x_k)} - \frac{1}{N}\Sigma_i(j-1)\sum_{k=1}^{N}\sum_{i=1}^{m} \frac{\alpha_i p_i(x_k)}{p(x_k)} \frac{(x_k-\mu_i(j-1))}{\Sigma_i(j-1)}}{\frac{1}{N}\sum_{k=1}^{N} \frac{p_i(x_k)}{p(x_k)}} \tag{7.75}$$

$$\Sigma_i(j) = \frac{\frac{1}{N}\sum_{k=1}^{N} \frac{p_i(x_k)}{p(x_k)}(x_k - \mu_i(j-1))^2}{\frac{1}{N}\sum_{k=1}^{N} \frac{p_i(x_k)}{p(x_k)}} \tag{7.76}$$

com $p_i(x_k)$ e $p(x_k)$ sendo funções de $\alpha_1(j-1), ..., \alpha_m(j-1), \mu_1(j-1), ...\mu_m(j-1), \Sigma_1(j-1), ..., \Sigma_m(j-1)$.

7.4 Procedimento de Peters e Walker modificado - restrição para variâncias

Vamos agora desenvolver as fórmulas de recorrência para obtenção dos parâmetros da mistura, considerando uma restrição para as variâncias. Assim, vamos procurar os parâmetros da mistura de forma que a relação abaixo seja satisfeita, isto é, todos os componentes possuem a mesma variância:

$$\Sigma_1 = \Sigma_2 = \cdots = \Sigma_m = \Sigma \tag{7.77}$$

Novamente, vamos considerar:

$$p(x) = \sum_{i=1}^{m} \alpha_i^0 p_i(x) \tag{7.78}$$

onde

$$x \in I\!R$$

$$i = 1, ..., m$$

$$\alpha_i^0 > 0 \tag{7.79}$$

$$\sum_{i=1}^{m} \alpha_i^0 = 1 \tag{7.80}$$

$$\Sigma_1 = \Sigma_2 = \cdots = \Sigma_m = \Sigma \tag{7.81}$$

$$p_i(x) = \frac{1}{\sqrt{2\pi}(\Sigma^0)^{1/2}} \exp\Big[-\frac{1}{2}(x-\mu_i^0)^2(\Sigma^0)^{-1}\Big] \tag{7.82}$$

A função de verossimilhança é dada por:

$$L(x_1,...,x_N) = \ln[f(x_1,...,x_N)] = \ln\Big[\prod_{k=1}^{N} p(x_k)\Big] = \sum_{k=1}^{N} \ln[p(x_k)] = \sum_{k=1}^{N} \ln\Big[\sum_{i=1}^{m} \alpha_i^0 p_i(x_k)\Big] \tag{7.83}$$

E a função de Lagrange:

$$F(\alpha_1,...,\alpha_m,\mu_1,...,\mu_m,\Sigma_1,...,\Sigma_m) = L + \lambda(\sum_{i=1}^{m} \alpha_i - 1) \tag{7.84}$$

Derivando F em relação aos parâmetros $(\alpha_i, \mu_i, \Sigma)_{i=1,...,m}$ e em relação a λ e igualando cada derivada parcial a zero, obtemos o ponto de máximo. Assim, derivando em relação a λ:

$$\frac{\partial F}{\partial \lambda} = 0 \Longrightarrow \sum_{i=1}^{m} \alpha_i - 1 = 0 \Longrightarrow \sum_{i=1}^{m} \alpha_i = 1 \tag{7.85}$$

Derivando em relação a α_i, temos:

$$\frac{\partial F}{\partial \alpha_i} = 0 \Longrightarrow \sum_{k=1}^{N} \frac{1}{\sum_{i=1}^{m} \alpha_i p_i(x_k)} p_i(x_k) + \lambda = 0 \tag{7.86}$$

Aplicando (7.78) em (7.86):

$$\sum_{k=1}^{N} \frac{p_i(x_k)}{p(x_k)} = -\lambda \tag{7.87}$$

A equação acima nos permite escrever:

$$\begin{cases} \alpha_1 \sum_{k=1}^{N} \frac{p_1(x_k)}{p(x_k)} = -\alpha_1\lambda \\ \vdots \\ \alpha_m \sum_{k=1}^{N} \frac{p_m(x_k)}{p(x_k)} = -\alpha_m\lambda \end{cases} \tag{7.88}$$

E portanto, somando as equações acima membro a membro:

$$\alpha_1 \sum_{k=1}^{N} \frac{p_1(x_k)}{p(x_k)} + ... + \alpha_m \sum_{k=1}^{N} \frac{p_m(x_k)}{p(x_k)} = -\alpha_1\lambda - ... - \alpha_m\lambda \tag{7.89}$$

E, rearranjando os termos

$$\sum_{k=1}^{N} \alpha_1 \frac{p_1(x_k)}{p(x_k)} + ... + \sum_{k=1}^{N} \alpha_m \frac{p_m(x_k)}{p(x_k)} = -(\alpha_1 + ... + \alpha_m)\lambda \tag{7.90}$$

Aplicando (7.80) na equação acima:

$$\sum_{k=1}^{N} \frac{\alpha_1 p_1(x_k) + ... + \alpha_m p_m(x_k)}{p(x_k)} = -\lambda \tag{7.91}$$

Utilizando (7.78), sabemos que $p(x) = \sum_{i=1}^{m} \alpha_i^0 p_i(x)$ então:

$$\sum_{k=1}^{N} \frac{p(x_k)}{p(x_k)} = -\lambda \tag{7.92}$$

E portanto:

$$N = -\lambda \tag{7.93}$$

Para encontramos uma fórmula de recorrência para α_i, vamos utilizar um artifício matemático multiplicando e dividindo a equação (10) por α_i e substituindo λ por $-N$. Assim, isolando α_i, temos:

$$\alpha_i = \frac{\alpha_i}{N} \sum_{k=1}^{N} \frac{p_i(x_k)}{p(x_k)} \tag{7.94}$$

Vamos agora obter uma fórmula de recorrência para μ_i, derivando F em relação a μ_i.

$$\frac{\partial F}{\partial \mu_i} = 0 \Longrightarrow \sum_{k=1}^{N} \frac{p_i(x_k)}{\sum_{i=1}^{m} \alpha_i p_i(x_k)} \Big[-\frac{1}{2} 2(x_k - \mu_i)(-1)\Sigma_i^{(-1)} \Big] = 0 \tag{7.95}$$

Por (7.78) e sabendo que $\Sigma_i > 0$:

$$\sum_{k=1}^{N} \frac{p_i(x_k)}{p(x_k)} \frac{(x_k - \mu_i)}{\Sigma_i} = 0 \Longrightarrow \sum_{k=1}^{N} x_k \frac{p_i(x_k)}{p(x_k)} - \sum_{k=1}^{N} \mu_i \frac{p_i(x_k)}{p(x_k)} = 0 \tag{7.96}$$

Isolando μ_i:

$$\mu_i = \frac{\sum_{k=1}^{N} x_k \frac{p_i(x_k)}{p(x_k)}}{\sum_{k=1}^{N} \frac{p_i(x_k)}{p(x_k)}} \tag{7.97}$$

Para fins de acuidade computacional, multiplica-se e divide-se o segundo membro da equação acima por $1/N$. Assim, obtemos a seguinte forma de recorrência para μ_i:

$$\mu_i = \frac{\frac{1}{N} \sum_{k=1}^{N} x_k \frac{p_i(x_k)}{p(x_k)}}{\frac{1}{N} \sum_{k=1}^{N} \frac{p_i(x_k)}{p(x_k)}} \tag{7.98}$$

Vamos agora derivar F em relação a Σ_i

$$\frac{\partial F}{\partial \Sigma_i} = 0 \Longrightarrow$$

$$\sum_{k=1}^{N} \frac{1}{\sum_{i=1}^{m} \alpha_i p_i(x_k)} \Big[\alpha_i \frac{1}{\sqrt{2\pi}} \Big(-\frac{1}{2} \Sigma_i^{-3/2} \exp(-\frac{1}{2}(x_k - \mu_i)^2 \Sigma_i^{-1}) \Big) +$$
$$\Big(\Sigma^{-1/2} \exp\Big(-\frac{1}{2}(x_k - \mu_i)^2 \Sigma_i^{-1} \Big) \Big(-\frac{1}{2}(x_k - \mu_i)^2 (-1) \Sigma^{-2} \Big) \Big) \Big] = 0 \tag{7.99}$$

Simplificando os termos, temos:

$$\frac{\partial F}{\partial \Sigma} = \sum_{k=1}^{N} \frac{1}{\sum_{i=1}^{m} \alpha_i p_i(x_k)} \sum_{i=1}^{m} \frac{\partial p_i}{\partial \Sigma} \tag{7.100}$$

onde

$$\frac{\partial p_i}{\partial \Sigma} = \frac{1}{\sqrt{2\pi}} \Big[-\frac{1}{2} \Sigma^{-\frac{3}{2}} \exp\Big(-\frac{1}{2}(x-\mu_i)^2 \Sigma^{-1} \Big) \Big]$$

$$\Big[\Sigma^{-\frac{1}{2}} \exp\Big(-\frac{1}{2}(x-\mu_i)^2 \Sigma^{-1} \Big) \Big(-\frac{1}{2} \Big) (x-\mu_i)^2 (-1) \Sigma^{-2} \Big]$$

$$= \frac{1}{\sqrt{2\pi}} \Sigma^{\frac{1}{2}} \exp\Big(-\frac{1}{2}(x-\mu_i)^2 \Sigma_i^{-1} \Big) \Big(-\frac{1}{2}\Sigma^{-1} + \frac{1}{2}\Sigma^{-2}(x-\mu_i)^2 \Big) =$$

$$p_i \Big(-\frac{1}{2}\Sigma^{-1} + \frac{1}{2}\Sigma^{-2}(x-\mu_i)^2 \Big)$$

então, fazendo $\frac{\partial F}{\partial \Sigma} = 0$:

$$\frac{\partial F}{\partial \Sigma} = \sum_{k=1}^{N} \frac{1}{p(x_k)} \sum_{i=1}^{m} \alpha_i p_i (-\frac{1}{2}\Sigma^{-1} + \frac{1}{2}\Sigma^{-2}(x-\mu_i)^2 \Big) = 0 \tag{7.101}$$

Reajustando a equação acima, temos:

$$\sum_{k=1}^{N} \frac{1}{p(x_k)} \Big[-\frac{1}{2\Sigma^2} \sum_{i=1}^{m} \alpha_i p_i \Big(\Sigma - (x-\mu_i)^2 \Big) \Big] \tag{7.102}$$

Como $\Sigma > 0$:

$$\sum_{k=1}^{N} \frac{1}{p(x_k)} \Big[\sum_{i=1}^{m} \alpha_i p_i \Sigma - \sum_{i=1}^{m} \alpha_i p_i (x-\mu_i)^2 \Big] = 0$$

Sendo Σ constante:

$$\sum_{k=1}^{N} \frac{1}{p(x_k)} \Big[\Sigma \sum_{i=1}^{m} \alpha_i p_i - \sum_{i=1}^{m} \alpha_i p_i (x-\mu_i)^2 \Big] = 0 \tag{7.103}$$

Por (7.78), sabemos que $\sum_{i=1}^{m} \alpha_i p_i = p(x_k)$. Temos:

$$\sum_{k=1}^{N} \frac{\Sigma p(x_k) - \sum_{i=1}^{m} \alpha_i p_i (x-\mu_i)^2}{p(x_k)} = 0 \tag{7.104}$$

Ou seja,

$$\sum_{k=1}^{N} \Sigma - \sum_{k=1}^{N} \frac{\sum_{i=1}^{m} \alpha_i p_i (x-\mu_i)^2}{p(x_k)} = 0 \tag{7.105}$$

Assim, isolando Σ, temos:

$$\Sigma = \frac{1}{N} \sum_{k=1}^{N} \frac{\sum_{i=1}^{m} \alpha_i p_i (x-\mu_i)^2}{p(x_k)} = 0 \tag{7.106}$$

As equações (7.94), (7.98) e (7.106), sugerem o seguinte o procedimento iterativo para a obtenção das estimativas dos parâmetros da mistura $\alpha_i, ..., \alpha_m, \mu_1, ..., \mu_m, \Sigma$. A partir de um conjunto de valores iniciais, também denominado chute inicial, $\alpha_i(0), ..., \alpha_m(0), \mu_1(0), ..., \mu_m(0), \Sigma(0)$, calculam-se os próximos conjuntos de parâmetros $\alpha_i(j), ..., \alpha_m(j), \mu_1(j), ..., \mu_m(j), \Sigma(j)$, com $j \in \mathbb{N}^*$, através das fórmulas de recorrência estabelecidas abaixo:

$$\alpha_i(j) = \frac{\alpha_i(j-1)}{N} \sum_{k=1}^{N} \frac{p_i(x_k)}{p(x_k)} \tag{7.107}$$

$$\mu_i(j) = \frac{\frac{1}{N}\sum_{k=1}^{N} x_k \frac{p_i(x_k)}{p(x_k)}}{\frac{1}{N}\sum_{k=1}^{N} \frac{p_i(x_k)}{p(x_k)}} \tag{7.108}$$

$$\Sigma_i(j) = \frac{1}{N} \sum_{k=1}^{N} \frac{\sum_{i=1}^{m} \alpha_i(j-1) p_i(x - \mu_i(j-1))^2}{p(x_k)} = 0 \tag{7.109}$$

com $p_i(x_k)$ e $p(x_k)$ sendo funções de $\alpha_1(j-1), ..., \alpha_m(j-1), \mu_1(j-1), ...\mu_m(j-1), \Sigma_1(j-1), ..., \Sigma_m(j-1)$.

7.5 Considerações sobre a convergência do algoritmo de Peters e Walker

Peters e Walker introduzem uma notação com o intuito de facilitar a representação dos elementos da mistura. Assim, vamos definir Θ em um espaço apropriado, tal que:

$$\Theta = \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_m \\ \mu_1 \\ \vdots \\ \mu_m \\ \Sigma_1 \\ \vdots \\ \Sigma_m \end{pmatrix} \tag{7.110}$$

Se estabelecermos:

$$A_i(\Theta) = \alpha_i$$
$$M_i(\Theta) = \mu_i$$
$$S_i(\Theta) = \Sigma_i$$

onde α_i, μ_i e $sigma_i$ são dadas pelas equações de recorrência identificadas em (7.26), (7.30) e (7.35), podemos escrever:

$$\Theta = \begin{pmatrix} A(\Theta) \\ M(\Theta) \\ S(\Theta) \end{pmatrix} \tag{7.111}$$

De forma mais genérica, podemos escrever Θ como uma função de $\epsilon \in I\!R$ arbitrário. Assim:

$$\Theta = \Phi_\epsilon(\Theta) = (1-\epsilon)\Theta + \epsilon \begin{pmatrix} A(\Theta) \\ M(\Theta) \\ S(\Theta) \end{pmatrix} \tag{7.112}$$

A equação sugere um procedimento iterativo para a obtenção das estimativas de máxima verossimilhança. Ou seja, tendo sido fornecido valores iniciais para Θ, obtém-se as estimativas dos parâmetros desejados através de iterações sucessivas que obedecem a relação :

$$\Theta^{(k+1)} = \Phi_\epsilon(\Theta^{(k)}) \tag{7.113}$$

onde $k = 1, 2, ...$ representa o número de iterações.

As fórmulas de recorrência desenvolvidas nas seções anteriores referem-se ao caso particular em que $\epsilon = 1$. Embora Peters e Walker tenham estudado formas de obtenção de um ϵ ótimo, isto é, que conduz a uma máxima taxa, uniforme e assintótica, de convergência local, em nosso trabalho aplicado, ϵ será uma variável que assumirá valores arbitrários determinados, não necessariamente um valor otimizado.

Apesar da perda de velocidade computacional, se escolhermos $0 < \epsilon < 2$, estaremos garantindo que o procedimento iterativo converge localmente para uma estimativa de máxima verossimilhança fortemente consistente.

7.6 Não-aplicabilidade do método de Peters e Walker

O processo de mistura de brownianos com a distribuição de Poisson (MP)representa um modelo mais elaborado do que o processo de mistura simples para a modelagem da dinâmica de flutuação do preço da ação. Porém, ao contrário do que foi visto nas seções anteriores para o processo de mistura simples, o procedimento de busca dos parâmetros através do método de Peters e Walker não pode ser aplicado no caso de mistura de brownianos com distribuição poissoniana dos pontos de alteração do comportamento do mercado. Nesta seção apresentaremos as razões desta não-aplicabilidade de forma indireta. Mostraremos que o método de Peters e Walker não se aplica ao processo MPA e os mesmos argumentos podem ser utilizados para concluir que o método também não funcionaria para o processo MP. Cabe lembrar que os processos MP e MPA foram discutidos no Capítulo 4.

Vamos supor inicialmente que os dados amostrais disponíveis refletem os valores dos preços do ativo-objeto, no caso ações, obtidos nas aberturas e fechamentos da Bolsa de Valores. A partir desta suposição, vamos verificar que o método de Peters e Walker não se aplica ao modelo MPA. É fácil verificar que a razão desta não-aplicabilidade será a mesma para qualquer conjunto de dados amostrais que sejam coletados em instantes fixados à priori, onde portanto, perde-se a informação do exato momento em que ocorre a mudança de comportamento do preço da ação.

Continuando nossa argumentação, consideremos que os cenários possíveis para o comportamento da ação sejam representados por $(\mu_1, \sigma), ..., (\mu_m, \sigma)$ e portanto, possuindo a mesma variância dos retornos. Vamos analisar o intervalo $[t_k, t_{k+1}]$. Suponhamos que este intervalo contenha um ponto do processo pontual de Poisson aproximado. Vamos designar por U a variável aleatória que determina a posição deste ponto no intervalo $[t_k, t_{k+1}]$, a partir do ponto t_k.

Então, dado que $U = u$ e que o preço da ação seguia um processo conforme o cenário i antes do ponto u e passou a seguir um processo representativo de um novo cenário j após o ponto u, temos:

$$S_{t_{k+1}} = S_{t_k} \exp(\mu_i u \sigma W_u) \exp(\mu_2(\Delta - u) + \sigma W_{\Delta - u}) \tag{7.114}$$

onde Δ é o comprimento do intervalo $[t_k, t_{k+1}]$.

Assim, o retorno do ativo $\ln\left(\frac{S_{t_{k+1}}}{S_{t_k}}\right)$ durante o intervalo $[t_k, t_{k+1}]$ tem distribuição equivalente à soma de duas normais, sendo uma com distribuição $\mathcal{N}(\mu_i(u), \sigma^2(u))$ e outra com distribuição $\mathcal{N}(\mu_j(\Delta - u), \sigma^2(\Delta - u))$.

Portanto, o retorno tem distribuição $\mathcal{N}(\mu_i u + \mu_j(\Delta - u), \sigma^2)$ dado que U assumiu o valor u e que a mudança de cenário no ponto u deu-se do i-ésimo para o j-ésimo cenário. Utilizando o fato de que U tem distribuição uniforme em $[t_k, t_{k+1}]$ teremos que, dado que ocorreu um ponto do prpcesso pontual de Poisson aproximado no intervalo $[t_k, t_{k+1}]$ e que neste ponto o cenário mudou do i-ésimo para o j-ésimo, a distribuição dos retornos durante este intervalo tem a seguinte função de densidade de probabilidades:

$$p_{ij}(x) = \frac{1}{\sqrt{2\pi}\sigma} \int_0^{\Delta} \frac{1}{\Delta} \exp\left(\frac{(x - \mu_i u - \mu_j(\Delta - u))^2}{2\sigma^2}\right) du \tag{7.115}$$

Suponhamos agora que desejamos inferir os valores dos parâmetros do processo MPA, isto é, os valores de $\lambda, \mu_1, ..., \mu_m, \sigma$, a partir dos dados amostrais. Os retornos

$$\ln\left(\frac{S_1}{S_0}\right), \ln\left(\frac{S_2}{S_1}\right), ..., \ln\left(\frac{S_{2N+1}}{S_{2N}}\right)$$

têm, então, a mesma distribuição. Note que estamos considerando os retornos entre a abertura e o fechamento de um mesmo pregão da Bolsa de Valores. Os retornos entre o fechamento de um pregão e a abertura do preção do dia seguinte em geral têm uma outra distribuição, tendo em vista a diferença entre os intervalos de amostragem. Para tentarmos aplicar o método de Peters e Walker só falta encontrar a densidade da distribuição de

$$\ln\left(\frac{S_{2N+1}}{S_{2N}}\right)$$

A partir da construção do processo pontual de Poisson aproximado, temos que a proporção de intervalos com um ponto é $\lambda\Delta$, enquanto a proporção de intervalos em que não há nenhum ponto é $1 - \lambda\Delta$, com $\Delta = |\mathcal{D}_{,}| = |\mathcal{D}_{\in}| = |\mathcal{D}_{\in\mathcal{N}}|$. Temos também que a probabilidade de o preço da ação seguir o processo relativo a um cenário i, num ponto qualquer, é α_i. Portanto, a distribuição de $\ln\left(\frac{S_{2N+1}}{S_{2N}}\right)$ tem a densidade

$$p(x) = (1 - \lambda\Delta) \sum_{i=1}^{m} \alpha_i p_i(x) + \lambda\Delta \sum_{i,j=1}^{m} \alpha_i \alpha_j p_{ij}(x) \tag{7.116}$$

onde

$$p_i(x) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{x - \mu_i)^2}{2\sigma^2}\right) \tag{7.117}$$

com $i = 1, \ldots, m$.

Fica fácil agora observar a razão que torna o método de Peters e Walker inaplicável para o caso acima:

1) $p_{ij}(x)$ não é densidade normal e não é uma mistura finita de densidades normais

2) os parâmetros $\alpha_{i}{}_{i=1}^{m}$ na expressão (7.116) encontram-se em uma forma não-linear.

Podemos notar que o segundo item (2) pode ser eliminado, porém, o primeiro item (1) é essencial, pois o método de Peters e Walker baseia-se fortemente na suposição de que $p(x)$ é uma mistura finita de densidades normais.

Bibliografia

[A] Arnold, L. *Stochastic differential equations: theory and applications.* John Wiley and Sons, 1974.

[BS] Black, F.; Scholes, M. *The pricing of options and corporate liabilities.* Journal of Political Economy 81, pp. 637-659, 1973.

[CRR] Cox, J.; Ross, s.; Rubinstein, M. *Option pricing: a simplified approach.* Journal of financial Economics 7, pp. 229-263, 1979.

[DP] Dixit, A. K.; Pyndick, R. S. *Investment under Uncertainty.* Princeton University Press, 1994.

[H] Hull, J. *Options, futures and other derivative securities.* Prentice Hall, 1993.

[M] Merton, R. *Theory of rational option pricing.* Bell Journal of Economics and Management Science 4, pp. 141-183, 1973.

[N] Neftci, S. *An Introduction to Mathematics of Financial Derivatives.* Academic Press, 1996.

[O] Oksendal B. *Introduction to Stochastic Calculus.* Springer-Verlag, 1992.

[PW1] Peters, C.; Walker, H. F. *An iterative procedure for obtaining maximum-likelihood estimates of the parameters for a mixture of normal distributions.* Siam Journal of Applied Mathematics 35, pp. 362-378, 1978.

[R] Ritchey, R. J. *Call option valuation for discrete normal mixtures.* The Journal of Financial Research 13, pp. 285-296, 1990.

[WHD] Wilmott, P.; Howison, S.; Dewynne, J. *The mathematics of financial derivatives: a student introduction.* Cambridge University Press, 1995.