**Universidade de São Paulo
Escola Superior de Agricultura “Luiz de Queiroz”**

Generalização do modelo AMMI para dados multiatributos

Jhessica Letícia Kirch

Dissertação apresentada para obtenção do título de
Mestra em Ciências. Área de concentração: Estatística e Experimentação Agronômica

**Piracicaba
2020**

Jhessica Letícia Kirch
Bacharela em Estatística

Generalização do modelo AMMI para dados multiatributos

versão revisada de acordo com a resolução CoPGr 6018 de 2011

Orientador:

Prof. Dr. **CÉSAR GONÇALVES DE LIMA**

Dissertação apresentada para obtenção do título de
Mestra em Ciências. Área de concentração: Estatística e Experimentação Agronômica

Piracicaba
2020

Dados Internacionais de Catalogação na Publicação
DIVISÃO DE BIBLIOTECA - DIBD/ESALQ/USP

Kirch, Jhessica Letícia

Generalização do modelo AMMI para dados multiatributos / Jhessica Letícia Kirch. – – versão revisada de acordo com a resolução CoPGr 6018 de 2011. – – Piracicaba, 2020 .

77 p.

Dissertação (Mestrado) – – USP / Escola Superior de Agricultura “Luiz de Queiroz”.

1. Estatística genética 2. Análise *multiway* 3. Dados de tripla entrada 4. Interação genótipos × ambientes × atributos . I. Título.

AGRADECIMENTOS

A Deus, autor e consumidor da minha fé, digno de toda honra, poder e majestade.

Aos meus pais pelo incentivo e apoio incondicional, sem os quais esse trabalho não seria possível.

Ao professor Dr. César Gonçalves de Lima pela orientação, dedicação, por compartilhar seus conhecimentos e experiências e pela confiança em mim depositada.

A todos os professores do Programa de Pós-Graduação em Estatística e Experimentação Agronômica (PPGEEA) que contribuíram para a minha formação acadêmica, em especial ao professor Dr. Carlos Tadeu dos Santos Dias, pela ajuda na elaboração e aperfeiçoamento desse trabalho.

Aos professores Dr. Sônia Maria De Stefano Piedade e Dr. Kuang Hongyu pelas contribuições no desenvolvimento dessa pesquisa.

A todos os funcionários do PPGEEA, sempre dispostos a ajudar e auxiliar a todos os alunos.

Aos colegas de curso pelo acolhimento e motivação durante toda a jornada. Aos colegas da Aliança Bíblica Universitária (ABU) pelo acolhimento e amizade.

Aos pesquisadores Dra. Acácia Mecejana Diniz Souza Spitti e professor Dr. Alisson Fernando Chiorato que forneceram os dados para esta dissertação.

E ao Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) pela concessão de bolsa de mestrado, fundamental para realização desta dissertação.

SUMÁRIO

Resumo	6
Abstract	7
Lista de Figuras	8
Lista de Tabelas	9
1 Introdução	11
2 Revisão bibliográfica	13
2.1 Interação genótipos $\times$ ambientes	13
2.2 Análise de variância (ANOVA)	15
2.3 Autovalores e autovetores de uma matriz	16
2.4 Decomposição por Valores Singulares (DVS)	17
2.5 Modelo AMMI	18
2.5.1 Ajuste do modelo AMMI	19
2.6 Gráfico <i>biplot</i>	21
2.7 Interação genótipos $\times$ ambientes para dados multiatributos	23
2.8 Metodologias para dados de três modos	24
2.9 Análise de variância para três fatores	26
2.10 Decomposição da interação tripla	27
2.11 Modelo Tucker3	28
2.11.1 Estimação do modelo Tucker3	29
2.11.2 Determinação da dimensionalidade de um modelo Tucker3	29
2.12 Gráfico <i>joint plot</i>	30
2.13 Análise de Procrustes	32
3 Material e Métodos	35
3.1 Material	35
3.2 Métodos	35
3.2.1 Análise de Variância para dois fatores	36
3.2.2 Modelo AMMI para um único atributo	39
3.2.3 Análise de variância para três fatores	39
3.2.4 Modelo Tucker3	44
3.2.5 Modelo AMMI para a interação tripla	45
3.2.6 Análise de Procrustes	46
3.2.7 <i>Software</i>	46
4 Resultados e Discussões	47
4.1 Análise de dados de dupla entrada	47
4.1.1 Modelo AMMI para o AT1	47
4.1.2 Modelo AMMI para o AT2	49
4.1.3 Modelo AMMI para o AT3	50
4.1.4 Modelo AMMI para o AT4	51
4.1.5 Modelo AMMI para o AT5	53
4.1.6 Modelo AMMI para o AT6	54
4.1.7 Modelo AMMI para o AT7	55
4.1.8 Modelo AMMI para o AT8	57
4.1.9 Conclusões gerais sobre os resultados das análises AMMI individuais	58
4.2 Análise de dados de tripla entrada	59

4.2.1	Ajuste do modelo Tucker3	59
4.2.2	Análise de Procrustes	62
4.3	Análise de dados de tripla entrada no caso bivariado	63
4.3.1	Ajuste do modelo Tucker3	63
4.3.2	Análise de Procrustes	66
5	Considerações Finais	67
	Referências	69
	Anexos	75
5.1	Correspondência por translação	75
5.2	Correspondência por dilatação	75
5.3	Correspondência por rotação	76

RESUMO

Generalização do modelo AMMI para dados multiatributos

Nos estudos de melhoramento de plantas, tem-se o objetivo de estudar a interação genótipos $\times$ ambientes, avaliando a estabilidade e adaptabilidade genotípica. O modelo aditivo com interação multiplicativa (AMMI) tem sido amplamente utilizado para casos em que se tem apenas uma variável resposta. O objetivo deste trabalho é revisar as metodologias utilizadas para análise da interação genótipos $\times$ ambientes em estudos de melhoramento genético em experimentos com multiatributos e indicar um método global para análise dos modelos AMMI multiatributo. Este método foi aplicado em um conjunto de dados relativo a um experimento de avaliação do escurecimento de grãos de feijão carioca pelo método de escurecimento natural do grão e pelo método de escurecimento acelerado instalado no delineamento experimental aleatorizado em blocos, no ano de 2016. Foram avaliados dezenove genótipos de feijoeiros do tipo carioca em seis ambientes no Estado de São Paulo e oito atributos. Os resultados indicaram que o modelo de Tucker3 é eficiente para lidar com o arranjo triplo da interação genótipos $\times$ ambientes $\times$ atributos e que a utilização do *joint plot* é aconselhável, pois permitiu avaliar a estabilidade e adaptabilidade genotípicas. O uso da análise de Procrustes associado ao modelo Tucker3 permitiu identificar conclusões globais para os diferentes atributos. Após a aplicação das metodologias, três genótipos foram destacados: o G15, que apresentou alto valor de luminosidade do grão e adaptação específica com o ambiente F; o G8, que apresentou alto valor de luminosidade do grão e adaptação específica com o ambiente A e o G12, genótipo com alta estabilidade e valor de luminosidade acima da média.

Palavras-chave: Interação genótipo $\times$ ambiente, Modelo Tucker3, Análise de Procrustes, *Multiway*

ABSTRACT

AMMI model generalization to multi-attribute data

In plant breeding studies, the objective is to study the genotype $\times$ environment interaction, evaluating genotypic stability and adaptability. The Additive Main-effects and Multiplicative Interaction model (AMMI) has been widely used for cases where there is only one variable response. The aim of this work is to review the methodologies used to analyze the interaction between genotypes and environments in genetic improvement studies in experiments with multi-attributes and to indicate a global method for analyzing the multi-attribute AMMI models. This method was applied to a data set relating to an experiment to evaluate the darkening of grains of carioca beans by natural darkening method and accelerated darkening method installed in a randomized block design in 2016. Nineteen carioca bean genotypes were evaluated in six environments in the State of São Paulo and eight attributes. The results indicated that the Tucker3 model is efficient to deal with the three-way array of genotypes $\times$ environments $\times$ attributes interaction and that the use of joint plot is advisable, as it allows to assess stability and genotypic adaptability. The use of Procrustes analysis associated with the Tucker3 model allows the identification of global conclusions for the different attributes. After the application of the methodologies, three genotypes were highlighted: the G15, which presented a high value of lightness of the grain and specific adaptation with the environment F; the G8, which presented a high value of lightness and specific adaptation with the environment A and G12, a genotype with high stability and above average lightness value.

Keywords: Genotypes $\times$ environments interaction, Tucker3 model, Procrustes analysis, Multiway

LISTA DE FIGURAS

2.1	Comportamento dos genótipos G1 e G2 nas condições ambientais A1 e A2. a) ausência de interação; b) interação simples; c) interação cruzada.	13
2.2	Comportamento dos genótipos G1 e G2 nas condições ambientais A1 e A2. a) ausência de efeito de genótipo com efeito de ambientes; b) ausência de efeito de ambientes com efeito de genótipo; c) ausência de efeito de genótipos e ambientes.	14
2.3	<i>Biplot</i> de um experimento com quatro genótipos em três ambientes a) $AMMI_1$; b) $AMMI_2$	22
2.4	Particionamento de um arranjo de três entradas em fatias horizontais, verticais e frontais.	27
2.5	Representação gráfica do modelo Tucker3.	29
2.6	<i>Joint plot</i> projetado dentro da primeira componente do terceiro modo considerando um experimento com três genótipos, três ambientes e dois tempos.	31
2.7	Passos para obtenção da estatística de Procrustes. a) figura inicial; b) translação; c) rotação; d) dilatação.	32
3.1	Diagrama de Hasse para os números de graus de liberdade associados aos termos do modelo (3.1)	37
3.2	Diagrama de Hasse para as somas de quadrados associadas aos termos do modelo (3.1)	38
3.3	Diagrama de Hasse para os números de graus de liberdade associados aos termos do modelo (3.3)	42
3.4	Diagrama de Hasse para as somas de quadrados associadas aos termos do modelo (3.3)	43
4.1	<i>Biplot</i> AMMI (CP_1 vs. Valores médios) para o AT1.	48
4.2	<i>Biplot</i> AMMI (CP_1 vs. CP_2) para o AT1.	48
4.3	<i>Biplot</i> AMMI (CP_1 vs. Valores médios) para o AT2.	49
4.4	<i>Biplot</i> AMMI (CP_1 vs. CP_2) para o AT2.	50
4.5	<i>Biplot</i> AMMI (CP_1 vs. Valores médios) para o AT3.	51
4.6	<i>Biplot</i> AMMI (CP_1 vs. CP_2) para o AT3.	51
4.7	<i>Biplot</i> AMMI (CP_1 vs. Valores médios) para o AT4.	52
4.8	<i>Biplot</i> AMMI (CP_1 vs. CP_2) para o AT4.	52
4.9	<i>Biplot</i> AMMI (CP_1 vs. Valores médios) para o AT5.	53
4.10	<i>Biplot</i> AMMI (CP_1 vs. CP_2) para o AT5.	54
4.11	<i>Biplot</i> AMMI (CP_1 vs. Valores médios) para o AT6.	55
4.12	<i>Biplot</i> AMMI (CP_1 vs. CP_2) para o AT6.	55
4.13	<i>Biplot</i> AMMI (CP_1 vs. Valores médios) para o AT7.	56
4.14	<i>Biplot</i> AMMI (CP_1 vs. CP_2) para o AT7.	56
4.15	<i>Biplot</i> AMMI (CP_1 vs. Valores médios) para o AT8.	57
4.16	<i>Biplot</i> AMMI (CP_1 vs. CP_2) para o AT8.	58
4.17	<i>Joint plot</i> projetado dentro da primeira componente do terceiro modo para o modelo Tucker3(2,2,1).	61
4.18	<i>Joint plot</i> projetado dentro da primeira componente do terceiro modo para o modelo Tucker3(4,4,1).	65

LISTA DE TABELAS

2.1	Esquema da análise de variância para g genótipos avaliados em e ambientes com r repetições.	15
2.2	Esquema da análise de variância para a atributos em g genótipos avaliados, e ambientes e r repetições.	26
3.1	Esquema da análise de variância para g genótipos avaliados em e ambientes com r blocos. .	38
3.2	Esquema da análise de variância para experimentos com a atributos em g genótipos avaliados e e ambientes com b blocos.	43
4.1	Soma de quadrados da interação genótipos $\times$ ambientes e <i>valor-p</i> do seu respectivo teste de hipótese na análise de variância para os atributos de escurecimento em grãos de feijoeiro. .	47
4.2	Decomposição das somas de quadrados da interação para o AT1 do escurecimento em grãos de feijoeiro.	47
4.3	Decomposição das somas de quadrados da interação para o AT2 do escurecimento em grãos de feijoeiro.	49
4.4	Decomposição das somas de quadrados da interação para o AT3 do escurecimento em grãos de feijoeiro.	50
4.5	Decomposição das somas de quadrados da interação para o AT4 do escurecimento em grãos de feijoeiro.	52
4.6	Decomposição das somas de quadrados da interação para o AT5 do escurecimento em grãos de feijoeiro.	53
4.7	Decomposição das somas de quadrados da interação para o AT6 do escurecimento em grãos de feijoeiro.	54
4.8	Decomposição das somas de quadrados da interação para o AT7 do escurecimento em grãos de feijoeiro.	56
4.9	Decomposição das somas de quadrados da interação para o AT8 do escurecimento em grãos de feijoeiro.	57
4.10	Resultados do procedimento de Timmerman-Kiers para selecionar o modelo de Tucker3 para o experimento sobre o escurecimento em grãos de feijoeiro.	59
4.11	Escores das componentes principais para um modelo de Tucker3(2,2,1) para o arranjo da interação tripla entre genótipos $\times$ ambientes $\times$ atributos.	60
4.12	Estatística de Procrustes (M^2).	63
4.13	Resultados do procedimento de Timmerman-Kiers para selecionar o modelo de Tucker3 para o experimento sobre o escurecimento em grãos de feijoeiro no caso bivariado.	63
4.14	Escores das componentes principais para o modelo de Tucker3(4,4,1) para as matrizes A , B e C	64
4.15	Escores das componentes principais para o arranjo da interação tripla do modelo de Tucker3(4,4,1).	64
4.16	Estatística de Procrustes (M^2) para o caso bivariado.	66

1 INTRODUÇÃO

As condições ambientais influenciam significativamente a produção agrícola, levando a variações consideráveis em seu rendimento e qualidade. Essa influência é denominada interação genótipos $\times$ ambientes (*Genotype-by-Environment Interaction* - GEI). Em estudos de melhoramento de plantas ou de culturas em múltiplos ambientes, tem-se o objetivo de estudar a GEI, avaliar a estabilidade e adaptabilidade genotípica, estabelecer relações entre genótipos, ambientes e genótipos $\times$ ambientes simultaneamente e prever a produção de certos genótipos em diferentes ambientes (Gauch, 2013).

Várias metodologias estatísticas têm sido propostas para o estudo da adaptabilidade e estabilidade genotípica. Entre essas ferramentas estão a regressão linear (Eberhart e Russell, 1966), o modelo AMMI (*Additive Main-effects and Multiplicative Interaction*) (Crossa et al., 1990; Gauch, 1992; Gauch e Zobel, 1996; Gauch, 2013), o modelo linear misto (Henderson, 1975; Smith et al., 2005; Resende, 2007), GGE biplot (*Genotype main effect + Genotype environment interaction*) (Yan e Kang, 2003; Yan e Holland, 2010) e modelo AMMI ponderado (Weighted AMMI) (Rodrigues et al., 2014).

Os modelos aditivos com interação multiplicativa (*Additive Main-effects and Multiplicative Interaction* - AMMI) têm sido amplamente utilizados na análise da interação GEI em experimentos balanceados quando se tem apenas uma variável resposta como, por exemplo, rendimento ou produtividade agrícola. Esse modelo se baseia na aplicação da decomposição por valores singulares, também chamada de análise de componentes principais (ACP), aos resíduos de um modelo de análise de variância (ANOVA) aditivo aplicado à tabela de médias da interação entre genótipos e ambientes (Rodrigues et al., 2014). A ACP reduz a dimensionalidade dos dados, gerando, a partir de combinações lineares das variáveis originais, novos eixos ou componentes principais, que retêm o máximo possível da variação original dos dados (Neisse e Hongyu, 2016).

Nos estudos de melhoramentos de grãos de feijoeiro, além do rendimento e resistência a patógenos, o escurecimento do tegumento dos grãos é uma característica de interesse que pode contribuir para a seleção genotípica, uma vez que grãos mais escuros têm menor valor comercial, pois o consumidor o associa a grãos antigos e com maior tempo de cocção. Essa associação não é correta, pois o escurecimento dos grãos também é influenciado pelo ambiente, época do ano e genótipo (Araújo et al., 2012).

De acordo com Spitti et al. (2019), a metodologia padrão para avaliação do escurecimento do grão utilizada atualmente é o escurecimento natural, que consiste na avaliação do valor de luminosidade de grãos acondicionados em embalagens plásticas e armazenados em condições de prateleira em temperatura e umidade ambiente, com aferição de coloração realizada a cada trinta dias. Métodos artificiais com iluminação ultravioleta têm sido utilizados como uma alternativa para diminuir o tempo necessário para avaliação do escurecimento dos grãos, como o método do escurecimento acelerado, no qual os grãos são expostos à luz ultravioleta e fluorescente e as aferições feitas a cada 24 horas.

Estudos envolvendo os métodos para avaliação do escurecimento do grão de diferentes genótipos de feijoeiro em múltiplos ambientes geram um conjunto de informações que precisam ser avaliadas simultaneamente, sendo necessário uma análise da interação entre genótipos e ambientes em uma metodologia multiatributo.

Algumas metodologias têm sido utilizadas para avaliar a GEI multiatributo utilizando a terminologia de dados de tripla entrada porque se aplicam a dados que podem ser classificados em três vias (genótipo, ambiente e atributo). Dentre elas destaca-se a análise de componentes principais de três modos (*three mode principal components analysis* - 3MPCA), algumas vezes denominada *Tucker3*, que permite a decomposição dos resíduos de um modelo de análise de variância com interação tripla, possibilitando uma interpretação similar àquela obtida pelo modelo AMMI.

Neste contexto, a presente dissertação tem os seguintes objetivos:

- I. Efetuar uma revisão das metodologias utilizadas para análise da interação entre genótipos e ambi-

entes em estudos de melhoramento genético em experimentos com multiatributos, essencialmente derivadas do modelo AMMI.

- II. Indicar um método global para análise dos modelos AMMI multiatributos como uma generalização do modelo AMMI usual para uma variável aleatória Y_{ijk} associada ao k -ésimo atributo no genótipo i e ambiente j .
- III. Aplicar o método indicado em dados de melhoramento de grãos de feijoeiro provenientes de um estudo GEI multiatributo para caracterização e comparação dos seus resultados.

2 REVISÃO BIBLIOGRÁFICA

2.1 Interação genótipos $\times$ ambientes

Nos estudos de melhoramento de plantas ou de culturas, é frequente a condução de experimentos em múltiplos ambientes, de forma a representar as diferentes condições climáticas da região ou país de interesse. A influência desses ambientes sobre rendimento ou qualidade dos genótipos em estudo é denominada interação genótipos $\times$ ambientes (*Genotype-by-Environment Interaction* - GEI).

Segundo Chaves (2001), essa interação deve ser encarada como um fenômeno biológico no melhoramento de plantas e não apenas como um simples efeito estatístico, uma vez que, em um experimento conduzido em múltiplos ambientes, serão beneficiados, em cada local, os indivíduos mais adaptados.

A adaptabilidade a diferentes condições ambientais foi uma característica de interesse desde o processo de domesticação de espécies, que provocou um aceleração nas mudanças evolutivas por meio das diferentes condições de cultivo e da seleção artificial provocada pelo homem. No melhoramento moderno, devido ao avanço tecnológico e ampliação do mercado agrícola, buscam-se cultivares com alto valor adaptativo aos mais variados ambientes, portanto a GEI, tomada como um fenômeno biológico natural, deve ser explorada e habilmente aproveitada no processo de seleção genotípica.

Se não ocorresse a interação GEI, um ensaio poderia ser realizado em um único ambiente fornecendo resultados universais. Além disso, se não ocorresse variação residual dentro de uma parcela ou entre parcelas de um mesmo experimento, não haveria necessidade de replicação. Assim, em um único ensaio seria possível identificar, sem erro, a variedade com valores superiores em todos os locais (Gauch e Zobel, 1996).

Kang (1993) destaca a importância dos ensaios multiambientais nos estágios iniciais do programa de melhoramento, uma vez que os genótipos descartados devido a testes limitados nas fases iniciais do experimento podem ter características superiores em outros locais ou épocas do ano. Além disso, os testes em vários ambientes permitem selecionar cultivares de desempenho mais consistente e geram melhores estimativas de componentes de variância.

A necessidade de repetição do experimento em vários ambientes gera um custo adicional aos programas de melhoramento, mas permite a identificação de cultivares com ampla adaptação bem como aqueles com adaptação específica a um determinado ambiente. A ampla adaptação fornece estabilidade diante de variabilidade ambiental, mas adaptações específicas fornecem uma vantagem de rendimento em ambientes particulares (Wade et al., 1999).

Figura 2.1. Comportamento dos genótipos G1 e G2 nas condições ambientais A1 e A2. a) ausência de interação; b) interação simples; c) interação cruzada.

Allard e Bradshaw (1964) ilustraram alguns tipos de interações em que são considerados dois

genótipos e dois ambientes. A Figura 2.1 apresenta três situações básicas que trazem diferentes consequências para o melhoramento.

Assume-se, na Figura 2.1, que Y é uma variável genérica e que o ambiente A2 é mais favorável em relação a Y. A situação a) representa a ausência de interação, uma vez que a mudança de ambiente afeta igualmente os genótipos 1 e 2, ou seja, a diferença entre os genótipos se mantém constante. Nas situações b) e c) a mudança de ambiente afeta cada genótipo de maneira diferente. A interação em b) é denominada simples ou quantitativa, pois em ambos os ambientes, o G1 apresenta valores de Y inferiores ao genótipo 2. No caso c), a interação é denominada cruzada ou qualitativa, pois há uma mudança na classificação dos genótipos.

Outros comportamentos podem ser observados considerando-se dois genótipos em duas condições ambientais, dentre os quais têm-se os três casos ilustrados na Figura 2.2. Na situação a) se considera que não existe diferença entre os genótipos, sendo este um caso específico da Figura 2.1 na situação a) em que a diferença entre genótipos é nula. Nas situações b) e c) da Figura 2.2 não existe efeito de ambientes, isto é, a mudança de ambiente não afeta os genótipos, contudo, em b), o genótipo 2 apresenta valores superiores aos de G1, enquanto que em c), não existe efeito de genótipos nem efeito de ambiente.

Figura 2.2. Comportamento dos genótipos G1 e G2 nas condições ambientais A1 e A2. a) ausência de efeito de genótipo com efeito de ambientes; b) ausência de efeito de ambientes com efeito de genótipo; c) ausência de efeito de genótipos e ambientes.

Os casos apresentados na Figura 2.2 são menos prováveis de ocorrer em experimentos de campo, em especial os casos b) e c), pois se esperam consideráveis variações dos fatores ambientais, como diferenças climáticas e tipos de solo, que geralmente resultam em resposta diferencial dos genótipos na variável de interesse. Quando essa diferença não ocorre, atribui-se aos genótipos um alto valor de adaptabilidade à variação ambiental.

Com exceção do caso c) da Figura 2.1, é possível basear-se na média dos ambientes para a seleção dos genótipos. Assim, nos casos em que existe uma diferença significativa entre os genótipos, seria inferido que o G2 é mais favorável à característica de interesse do que o genótipo 1. Na situação c) da Figura 2.1, a seleção baseada na média não é suficiente, podendo levar a uma seleção incorreta de genótipos no melhoramento (Chaves, 2001).

Nas três situações ilustradas na Figura 2.1 e na situação a) da Figura 2.2, assume-se o caso em que o ambiente A2 é mais favorável a Y do que o ambiente A1 e nas três situações na Figura 2.1 e na situação b) da Figura 2.2, assume-se o caso em que o G2 é mais favorável a Y do que o genótipo G1, entretanto, para todos os casos os genótipos e os ambientes podem ser permutados, obtendo-se interpretações semelhantes. Pode-se, ainda, considerar um número superior de genótipos e/ou de ambientes em todas as situações ilustradas, gerando uma relação complexa que demanda métodos adequados de análise.

Desde a década de 70 várias metodologias foram utilizadas na tentativa de capturar conjuntamente os efeitos de genótipos e a interação genótipos $\times$ ambientes (Kang, 2001). Na pesquisa genética, a

metodologia AMMI (*Additive Main-effects and Multiplicative Interaction*) tem sido amplamente utilizada por possibilitar a identificação de genótipos de alta produtividade e largamente adaptados, bem como na identificação de zoneamento agrícola e recomendação regionalizada. Além disso, os resultados podem ser representados em um gráfico *biplot* que mostra efeitos principais e de interação para os genótipos e os ambientes (Gauch e Zobel, 1996).

2.2 Análise de variância (ANOVA)

Considerando um estudo da GEI em que g genótipos ($i = 1, 2, \dots, g$) foram testados em e ambientes ($j = 1, 2, \dots, e$) em um experimento com r repetições ($k = 1, 2, \dots, r$), o modelo clássico (aditivo) da análise de variância (ANOVA) pode ser definido como (Perkins e Jinks, 1968):

$$y_{ijk} = \mu + g_i + e_j + (ge)_{ij} + \varepsilon_{ijk} \quad (2.1)$$

em que

y_{ijk} é o valor observado do i -ésimo genótipo no j -ésimo ambiente e na k -ésima repetição;

μ é uma constante comum às respostas;

g_i é o efeito fixo do genótipo i ($i = 1, 2, \dots, g$);

e_j é o efeito fixo do ambiente j ($j = 1, 2, \dots, e$);

$(ge)_{ij}$ é o efeito da interação do i -ésimo genótipo com o j -ésimo ambiente;

ε_{ijk} é o erro experimental associado à observação, assumindo $\varepsilon_{ijk} \sim N(0, \sigma^2)$ e $cov(\varepsilon_{ijk}, \varepsilon_{i'j'k'}) = 0$ para $(i, j, k) \neq (i', j', k')$.

A Tabela 2.1 apresenta o esquema do quadro da análise conjunta de variância do experimento, assumindo um delineamento experimental inteiramente aleatorizado. O número de graus de liberdade do resíduo foi obtido pela diferença do número de graus de liberdade total e a soma dos número de graus de liberdade dos genótipos, ambientes e da interação $G \times E$.

Tabela 2.1. Esquema da análise de variância para g genótipos avaliados em e ambientes com r repetições.

Fonte de variação	GL
Genótipos (G)	$(g - 1)$
Ambientes (E)	$(e - 1)$
Interação (G×E)	$(g - 1)(e - 1)$
Resíduo	$ge(r - 1)$
Total	$ger - 1$

É possível encontrar os estimadores dos parâmetros da equação (2.1) pelo método de mínimos quadrados. Admitindo-se as condições marginais $\sum_{i=1}^g g_i = \sum_{j=1}^e e_j = \sum_{i=1}^g \sum_{j=1}^e (ge)_{ij} = 0$, as hipóteses testadas no modelo são:

- i. $H_0 : g_i = 0$, para todo i
 $H_1 : g_i \neq 0$, para algum i
- ii. $H_0 : e_j = 0$, para todo j
 $H_1 : e_j \neq 0$, para algum j

- iii. $H_0 : (ge)_{ij} = 0$, para todo i, j
 $H_1 : (ge)_{ij} \neq 0$, para algum i, j

Para testar cada hipótese H_0 , utiliza-se o teste F, dado pela razão entre o quadrado médio do efeito a ser testado e o quadrado médio do resíduo. Ao rejeitar a hipótese nula em (i), conclui-se que existem evidências de que os efeitos de genótipos não são nulos, dado o nível de significância do teste. A mesma interpretação pode ser feita em (ii) para o efeito dos ambientes. Em (iii) testa-se a existência de efeito diferencial do genótipo em pelo menos um ambiente ou do efeito diferencial dos ambientes em pelo menos um dos genótipos. Observa-se que, quando a hipótese nula em (iii) é rejeitada, deve-se decompor a interação $G \times E$ e testar o efeito do genótipo em cada ambiente e o efeito do ambiente em cada genótipo, desconsiderando os resultados dos testes (i) e (ii) sobre os efeitos principais de genótipo de ambiente.

Ressalta-se que esses testes só são válidos se os pressupostos assumidos para os erros do modelo estiverem satisfeitos. Nos casos em que pressuposições da ANOVA não são satisfeitas, a família de transformações Box-Cox (Box e Cox, 1964) tem sido amplamente utilizada por produzir, com base na utilização de um valor λ que maximize o estimador de máxima verossimilhança e minimize o resíduo, uma transformação da variável resposta tal que as pressuposições de homocedasticidade e normalidade dos resíduos sejam simultaneamente satisfeitas (Azevedo et al., 2016).

2.3 Autovalores e autovetores de uma matriz

Para cada matriz quadrada $\mathbf{A}_{(p \times p)}$, existe um escalar λ , denominado valor singular ou autovalor de $\mathbf{A}$, e um vetor $\mathbf{x}$ não nulo, denominado autovetor ou vetor característico, tal que $\mathbf{Ax} = \lambda\mathbf{x}$. Para calcular λ e o autovetor $\mathbf{x}$ de $\mathbf{A}$ associado a λ , resolve-se o sistema:

$$(\mathbf{A} - \lambda\mathbf{I})\mathbf{x} = \mathbf{0} \quad (2.2)$$

em que $\mathbf{0}$ é um vetor nulo de dimensão $p \times 1$ e $\mathbf{I}$ é a matriz identidade de ordem p . O sistema (2.2) admite solução não trivial se, e somente se, o determinante da matriz quadrada $(\mathbf{A} - \lambda\mathbf{I})$ for igual a zero, isto é,

$$|\mathbf{A} - \lambda\mathbf{I}| = 0 \quad (2.3)$$

Se $\mathbf{A}$ for de dimensão $p \times p$, então a equação característica (2.3) terá p raízes, isto é, $\mathbf{A}$ terá p autovalores $(\lambda_1, \lambda_2, \dots, \lambda_p)$, que não são necessariamente todos distintos, diferentes de zero ou todos reais (Rencher e Schaafje, 2007).

Assim, se $p = 2$, então

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$$

Por (2.3), temos

$$|\mathbf{A} - \lambda\mathbf{I}| = 0$$

$$\left| \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} - \lambda \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \right| = 0$$

$$\left| \begin{bmatrix} a_{11} - \lambda & a_{12} \\ a_{21} & a_{22} - \lambda \end{bmatrix} \right| = 0$$

$$\begin{aligned}
(a_{11} - \lambda)(a_{22} - \lambda) - a_{12}a_{21} &= 0 \\
a_{11}a_{22} - a_{11}\lambda - a_{22}\lambda + \lambda^2 - a_{12}a_{21} &= 0 \\
\lambda^2 - (a_{11} + a_{22})\lambda + (a_{11}a_{22} - a_{12}a_{21}) &= 0
\end{aligned} \tag{2.4}$$

As raízes λ_1 e λ_2 da equação (2.4) devem ser obtidas algebricamente. Observa-se que $\lambda_1 + \lambda_2 = a_{11} + a_{22}$ e $\lambda_1 * \lambda_2 = \det(\mathbf{A})$. Os autovetores associados a cada um dos autovalores λ_1 e λ_2 são dados pelo sistema (2.2). Para λ_1 , temos

$$\begin{aligned}
(\mathbf{A} - \lambda_1 \mathbf{I})\mathbf{x}_1 &= \mathbf{0} \\
\left(\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} - \lambda_1 \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \right) \mathbf{x}_1 &= \begin{bmatrix} 0 \\ 0 \end{bmatrix} \\
\begin{bmatrix} a_{11} - \lambda_1 & a_{12} \\ a_{21} & a_{22} - \lambda_1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} &= \begin{bmatrix} 0 \\ 0 \end{bmatrix} \\
\begin{cases} (a_{11} - \lambda_1)x_1 + a_{12}x_2 = 0 \\ a_{21}x_1 + (a_{22} - \lambda_1)x_2 = 0 \end{cases} & \tag{2.5}
\end{aligned}$$

Resolvendo-se as equações em (2.5), serão obtidos os valores de x_1 e x_2 do vetor $\mathbf{x}_1$ associada ao autovalor λ_1 . O autovetor associado ao autovalor λ_2 será obtido de maneira análoga.

2.4 Decomposição por Valores Singulares (DVS)

A decomposição por valores singulares (DVS), proposta por Eckart e Young (1936), é uma técnica da álgebra matricial com propriedades matemáticas fortes que representa a base teórica para diversas técnicas, como a Análise de Componentes Principais, uma vez que a DVS é uma generalização da decomposição espectral, restrita a matrizes quadradas e simétricas.

Por meio da decomposição por valores singulares é possível obter uma aproximação para uma matriz real qualquer, de posto p , em p matrizes mutuamente ortogonais e de posto unitário (Souza, 1988). Assim, a DVS para uma matriz $\mathbf{B}_{(l \times c)}$ é dada por

$$\mathbf{B} = \sum_k \lambda_k \boldsymbol{\gamma}_k \boldsymbol{\alpha}_k^T \tag{2.6}$$

em que λ_k é o k -ésimo autovalor de $\mathbf{B}$ e $\boldsymbol{\gamma}_k$ e $\boldsymbol{\alpha}_k^T$ são autovetores singulares (linha e coluna, respectivamente) associados ($k = 1, 2, \dots, p$; com $p \leq \min(l, c)$). O sobrescrito T em $\mathbf{y}^T$ indica o vetor $\mathbf{y}$ transposto. A melhor aproximação de posto n para a matriz $\mathbf{B}$ é obtida pelo somatório de um número n de termos, inferior a p .

Os p valores singulares são dispostos, em ordem decrescente, em uma matriz diagonal $\boldsymbol{\Lambda}_p$ e os respectivos autovetores singulares $\boldsymbol{\gamma}_k$ e $\boldsymbol{\alpha}_k^T$, em duas matrizes $\mathbf{U}_{(l \times p)}$ e $\mathbf{V}_{(p \times c)}^T$, respectivamente. Assim, a expressão (2.6) pode ser escrita da forma

$$\mathbf{B} = \mathbf{U}\mathbf{\Lambda}\mathbf{V}^T \quad (2.7)$$

de forma que $\mathbf{U}$ tem os p vetores $\boldsymbol{\gamma}_k$ em suas colunas e $\mathbf{V}^T$ os p vetores $\boldsymbol{\alpha}_k^T$ em suas linhas.

Por exemplo, se considerarmos uma matriz $\mathbf{B}_{(2 \times 2)}$, deve-se obter, inicialmente os $p = 2$ valores singulares e seus respectivos autovetores $\boldsymbol{\gamma}_k$ e $\boldsymbol{\alpha}_k^T$ ($k = 2$) denominados autovetores singulares coluna e linha, respectivamente. É possível obter os valores singulares de uma matriz por meio da raiz quadrada dos autovalores (λ_k^2) de uma das correspondentes matrizes simétricas $\mathbf{B}\mathbf{B}^T$ e $\mathbf{B}^T\mathbf{B}$, que possuem os mesmos autovalores (Duarte e Vencovsky, 1999). Os autovetores de $\mathbf{B}\mathbf{B}^T$ associados a cada λ_k^2 , dados em ordem decrescente destes autovalores, correspondem aos p vetores-coluna que formam a matriz $\mathbf{U}$, isto é, os vetores singulares-coluna de $\mathbf{B}$. De maneira equivalente, os p autovetores de $\mathbf{B}^T\mathbf{B}$ são os vetores singulares-linha de $\mathbf{B}$, dados como linhas de $\mathbf{V}^T$.

Toma-se como exemplo a matriz $\mathbf{B}$ dada por

$$\mathbf{B} = \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix}$$

As matrizes $\mathbf{B}\mathbf{B}^T$ e $\mathbf{B}^T\mathbf{B}$ são obtidas por multiplicação usual de matrizes e os autovalores de $\mathbf{B}\mathbf{B}^T$ e $\mathbf{B}^T\mathbf{B}$ e os respectivos autovetores associados a cada autovalor de cada matriz foram obtidos conforme indicado em (2.2). Para a matriz $\mathbf{B}\mathbf{B}^T$, têm-se os autovalores λ_1^2 e λ_2^2 , com $\lambda_1^2 > \lambda_2^2$, e os autovetores associados a eles são, respectivamente,

$$\boldsymbol{\gamma}_1 = \begin{bmatrix} \gamma_{111} \\ \gamma_{121} \end{bmatrix} \quad \boldsymbol{\gamma}_2 = \begin{bmatrix} \gamma_{211} \\ \gamma_{221} \end{bmatrix}$$

com $\gamma_{111}^2 + \gamma_{121}^2 = 1$ e $\gamma_{211}^2 + \gamma_{221}^2 = 1$. Para a matriz $\mathbf{B}^T\mathbf{B}$, tem-se os mesmos autovalores, porém os autovetores associados a λ_1^2 e λ_2^2 são, respectivamente,

$$\boldsymbol{\alpha}_1 = \begin{bmatrix} \alpha_{111} \\ \alpha_{121} \end{bmatrix} \quad \boldsymbol{\alpha}_2 = \begin{bmatrix} \alpha_{211} \\ \alpha_{221} \end{bmatrix}$$

com $\alpha_{111}^2 + \alpha_{121}^2 = 1$ e $\alpha_{211}^2 + \alpha_{221}^2 = 1$.

A decomposição por valores singulares de $\mathbf{B}$, dada por (2.6) será

$$\mathbf{B} = \sum_k \lambda_k \boldsymbol{\gamma}_k \boldsymbol{\alpha}_k^T = \lambda_1 \boldsymbol{\gamma}_1 \boldsymbol{\alpha}_1^T + \lambda_2 \boldsymbol{\gamma}_2 \boldsymbol{\alpha}_2^T$$

em que λ_k é a raiz quadrada de λ_k^2 . De maneira equivalente, por (2.7), tem-se que a decomposição por valores singulares de $\mathbf{B}$ é

$$\mathbf{B} = \mathbf{U}\mathbf{\Lambda}\mathbf{V}^T = \begin{bmatrix} \gamma_{111} & \gamma_{211} \\ \gamma_{121} & \gamma_{221} \end{bmatrix} \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix} \begin{bmatrix} \alpha_{111} & \alpha_{211} \\ \alpha_{121} & \alpha_{221} \end{bmatrix}$$

2.5 Modelo AMMI

O modelo de efeitos principais aditivos e interação multiplicativa (AMMI) combina a análise de variância (ANOVA) e a análise de componentes principais (ACP) em um único modelo. Inicialmente

é aplicada a ANOVA para particionar a variação total em genótipo, ambiente, interação entre genótipo e ambiente e na variação não explicada (resíduo) e depois é aplicada a análise de componentes principais à interação GEI (Gauch, 2013). Assim, tanto os genótipos como a interação genótipos $\times$ ambientes são analisadas, mas separadamente, gerando um conjunto de informações fundamentais aos melhoristas.

A análise de componentes principais (ACP) foi desenvolvida por Pearson (1901) e a análise de variância (ANOVA) por Fisher (1918). De acordo com Gauch (1992), Fisher e Mackenzie (1923) foram os primeiros a aplicar ambas as análises (ANOVA e ACP), separadamente, em um conjunto de dados proveniente de um ensaio que avaliou o rendimento de 12 variedades de batata a partir da aplicação de 6 fertilizantes (ambientes), com três repetições. Entretanto, eles aplicaram a ACP nos dados originais e não na matriz com as médias da interação, como ocorre na análise AMMI. Metodologia semelhante à AMMI foi aplicada por Williams (1952) e Pike e Silverberg (1952) analisando um experimento químico. Os estudos da metodologia AMMI foram retomados por Gollob (1968) e Mandel (1971), coincidindo com o avanço tecnológico dos computadores digitais. A aplicação do modelo AMMI à pesquisa agrícola foi difundida por Kempton (1984) e Zobel et al. (1988), mas, de acordo com Annicchiarico (2002), seu uso se propagou graças ao trabalho de Gauch (1992).

2.5.1 Ajuste do modelo AMMI

No ajuste do modelo de efeitos principais aditivos e interação multiplicativa (AMMI) os efeitos principais, isto é, a média, os efeitos genotípicos e ambientais e a interação GEI são estimados por uma análise de variância comum (univariada) e a matriz de médias da interação genótipos $\times$ ambientes será analisada pela decomposição por valores singulares ou por análise de componentes principais, como grande número de autores preferem referir (Duarte e Vencovsky, 1999). Assim, a resposta média sobre r repetições ou blocos do i -ésimo genótipo no j -ésimo ambiente é dada por meio do modelo:

$$y_{ij} = \mu + g_i + e_j + \delta_{ij} + \varepsilon_{ij} \quad (2.8)$$

em que

y_{ij} corresponde a resposta média do i -ésimo genótipo no j -ésimo ambiente;

μ é uma constante comum;

g_i e e_j correspondem, respectivamente, ao efeito principal do genótipo i ($i = 1, 2, \dots, g$) e do ambiente j ($j = 1, 2, \dots, e$);

δ_{ij} representa o resíduo da não aditividade dos efeitos principais, equivalente ao termo $(ge)_{ij}$ do modelo (2.1);

ε_{ij} é o erro experimental associado à y_{ij} .

Por meio da decomposição por valores singulares (ou por análise de componentes principais) é possível modelar uma estrutura para os termos δ_{ij} do modelo (2.8) em um conjunto de componentes multiplicativos mais um erro residual

$$\delta_{ij} = \sum_{k=1}^n \lambda_k \alpha_{ik} \gamma_{jk} + \rho_{ij} \quad (2.9)$$

em que λ_k é a raiz quadrada do k -ésimo autovalor das matrizes $(\mathbf{GE})(\mathbf{GE})^T$ e $(\mathbf{GE})^T(\mathbf{GE})$, com $(\mathbf{GE}) = [(\hat{ge})_{ij}]$ obtida pela estimativa dos efeitos principais para o modelo (2.8) e $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_p$, α_{ik} é o i -ésimo valor estimado do vetor de escores genotípicos $\boldsymbol{\alpha}_k$ associado a λ_k , γ_{jk} é o j -ésimo valor estimado do vetor de escores ambientais $\boldsymbol{\gamma}_k$ associado a λ_k e ρ_{ij} é o ruído.

Na matriz de interações $(\mathbf{GE}) = [(\hat{ge})_{ij}]$, dada por

$$(\mathbf{GE}) = \begin{bmatrix} (\hat{ge})_{11} & (\hat{ge})_{12} & \cdots & (\hat{ge})_{1e} \\ (\hat{ge})_{21} & (\hat{ge})_{22} & \cdots & (\hat{ge})_{2e} \\ \vdots & \vdots & \ddots & \vdots \\ (\hat{ge})_{g1} & (\hat{ge})_{g2} & \cdots & (\hat{ge})_{ge} \end{bmatrix} \quad (2.10)$$

cada elemento $(\hat{ge})_{ij}$ é obtido por meio da relação

$$(\hat{ge})_{ij} = Y_{ij} - \bar{Y}_{i.} - \bar{Y}_{.j} + \bar{Y}_{..} \quad (2.11)$$

em que

Y_{ij} corresponde a resposta média do i -ésimo genótipo no j -ésimo ambiente;

$\bar{Y}_{i.}$ é a média do genótipo i ;

$\bar{Y}_{.j}$ é a média do ambiente j ;

$\bar{Y}_{..}$ é a média geral do experimento.

Observa-se que a DVS permite particionar toda a soma de quadrados da matriz $(\mathbf{GE}) = [(\hat{ge})_{ij}]$. Como $(\mathbf{GE})$ é uma matriz de desvios centrados (tem média nula), a soma de quadrados de seus elementos equivale a soma de quadrados da interação, portanto a $SQ_{G \times E}$ é desdobrada em componentes relativos a cada termo da DVS, ou cada eixo de interação da ACP, denominado componente principal (CP). A soma de quadrados relativa ao k -ésimo termo (ou eixo) de interação equivale ao quadrado do valor singular correspondente, então a $SQ_{G \times E}$ para a componente principal k é igual a λ_k^2 (Duarte e Vencovsky, 1999).

O índice k pode assumir valores de 1 a p com $p \leq \min(g-1, e-1)$. Tomando-se o somatório até n com $n < p$, determina-se uma aproximação de mínimos quadrados para a matriz $(\mathbf{GE})$ pelos n primeiros termos da DVS (ou n primeiros componentes principais), deixando um resíduo adicional ρ_{ij} (Gabriel, 1978). Se $n = p$, tem-se a decomposição exata da matriz, implicando em um resíduo adicional ρ_{ij} nulo. Sendo, portanto, possível escrever os componentes aditivos e multiplicativos do modelo AMMI da seguinte forma:

$$y_{ij} = \mu + g_i + e_j + \sum_{k=1}^p \lambda_k \alpha_{ik} \gamma_{jk} + \varepsilon_{ij} \quad (2.12)$$

em que $\sum_{k=1}^p \lambda_k \alpha_{ik} \gamma_{jk} = \sum_{k=1}^n \lambda_k \alpha_{ik} \gamma_{jk} + \sum_{k=n+1}^p \lambda_k \alpha_{ik} \gamma_{jk}$ com $n < p$. Este modelo (2.12) pode ser reescrito como:

$$y_{ij} = \mu + g_i + e_j + \sum_{k=1}^n \lambda_k \alpha_{ik} \gamma_{jk} + \rho_{ij} + \varepsilon_{ij}$$

em que n é o número de eixos ou componentes principais da interação $G \times E$ a serem considerados no modelo (Gauch e Zobel, 1996). Dependendo do número de eixos da interação retidos, o modelo é denotado $AMMI_0, AMMI_1, \dots, AMMI_n$, constituindo uma família de modelos AMMI. Em $AMMI_0$ nenhum eixo é ajustado (denominando modelo aditivo sem interação), em $AMMI_1$ apenas o primeiro eixo da interação é utilizado. Se $n = p$, isto é, se todos os eixos da interação forem retidos, o modelo apresentará um ajuste perfeito à matriz de médias. Entretanto, o objetivo do modelo AMMI é resumir grande parte da interação genótipos $\times$ ambientes em poucos eixos, resultando na seleção de um modelo AMMI reduzido que descarta o resíduo adicional (Duarte e Vencovsky, 1999).

A seleção do número de eixos que explicam o comportamento da interação é determinado considerando a proporção da $SQ_{G \times E}$ acumulada até o eixo n ($\sum_{k=1}^n \lambda_k^2 / SQ_{G \times E}$), baseando-se na significância

do teste F , dado pela razão entre o quadrado médio da componente principal k (QM_{CP_k}) e o quadrado médio do resíduo. O método de Gollob é um dos métodos utilizados para determinar os números de graus de liberdade para o modelo AMMI, sendo calculado como $GL_{CP_k} = g + e - 1 - 2k$, com $k = 1, 2, \dots, p$, $p = \min[(g - 1), (e - 1)]$ e CP_k sendo o k -ésimo eixo das componentes principais.

Deve-se ressaltar que a metodologia AMMI, mesmo sendo uma técnica multivariada, apresenta uma característica importante e incomum das técnicas multivariada: a análise AMMI envolve apenas uma variável resposta (como produtividade, usualmente) e é essa única característica que é submetida a uma técnica estatística multivariada.

2.6 Gráfico *biplot*

O *biplot* é uma representação gráfica baseada na decomposição por valores singulares de uma matriz numérica de dados. O conceito do *biplot* é atribuído à Gabriel (1971) e pode ser utilizado para representar em um único gráfico os efeitos de interação entre todos os genótipos e ambientes do modelo AMMI (Duarte e Vencovsky, 1999).

Dada a matriz $\mathbf{B}_{(l \times c)}$ de posto $p = \min(l, c)$ e seja a sua DVS como dada na equação (2.7) ($\mathbf{B} = \mathbf{U}\mathbf{\Lambda}\mathbf{V}^T$), de forma que os p valores singulares são dispostos, em ordem decrescente, na matriz diagonal $\mathbf{\Lambda}_p$ e que $\mathbf{U}_{(l \times p)}$ possua os p vetores γ_k em suas colunas e $\mathbf{V}_{(p \times e)}^T$ os p vetores α_k^T em suas linhas. A decomposição da equação (2.7) pode ser reescrita como

$$\mathbf{B} = \mathbf{U}(\mathbf{\Lambda}^{1/2}\mathbf{\Lambda}^{1/2})\mathbf{V}^T = (\mathbf{U}\mathbf{\Lambda}^{1/2})(\mathbf{\Lambda}^{1/2}\mathbf{V}^T)$$

Denotando $\mathbf{G} = (\mathbf{U}\mathbf{\Lambda}^{1/2})$ e $\mathbf{H}^T = (\mathbf{\Lambda}^{1/2}\mathbf{V}^T)$, isto é, $\mathbf{B} = \mathbf{G}\mathbf{H}^T$, em que $\mathbf{G}$ é a matriz $l \times p$ de l vetores de p componentes, em que os elementos correspondem às coordenadas para as l linhas e $\mathbf{H}^T$ é a matriz $p \times c$ com c vetores de p componentes, que contém as coordenadas das c colunas (Gabriel, 1971). Dessa forma, o *biplot* é capaz de representar simultaneamente as linhas e as colunas da matriz $\mathbf{B}$.

No modelo AMMI, a aproximação DVS é aplicada a matriz de interações ($\mathbf{GE}$) da equação (2.2), de forma que $(\hat{\mathbf{G}}\mathbf{E}) = (\mathbf{U}\mathbf{\Lambda}^{1/2})(\mathbf{\Lambda}^{1/2}\mathbf{V}^T)$, em que a matriz $\mathbf{G}_{(g \times n)}$ é composta pelos g vetores com n componentes, correspondendo as coordenadas dos genótipos para cada eixo singular selecionado e a matriz $\mathbf{H}^T$ é composta pelos e vetores com n componentes, que corresponde às coordenadas de ambientes para os mesmos n eixos (Duarte e Vencovsky, 1999).

Na prática, é possível representar todos os eixos retidos na análise AMMI sempre aos pares. A Figura 2.3 apresenta dois exemplos de *biplot*, o primeiro, denominado $AMMI_1$ contém no eixo das abcissas as médias da variável resposta do estudo e no eixo das ordenadas, a primeira componente principal (CP_1). O segundo *biplot* da Figura 2.3 apresenta no eixo das abcissas a CP_1 e no eixo das ordenadas, a segunda componente principal (CP_2). Se mais eixos forem retidos na análise, é possível construir tantos *biplots* quanto forem necessários de forma que o eixo das abcissas represente a CP_k e o eixo das ordenadas a CP_{k+1} , entretanto os primeiros eixos são invariavelmente os mais representativos e, como resultado disso, de interpretação determinante.

A interpretação do *biplot* está condicionada aos eixos retratados e o percentual de variação explicada por eles, que geralmente é disposto na figura. O $AMMI_1$ retrata a resposta média no eixo das abcissas, assim, genótipos e ambientes a direita no gráfico apresentam, em média, maiores valores da variável resposta; genótipos e ambientes a esquerda no gráfico apresentam, em média, menores valores da variável resposta. No caso a) da Figura 2.3, identifica-se que o ambiente 1 (A1) apresenta a maior média de resposta e o A2, a menor, sendo possível caracterizar estes ambientes em relação a sua resposta média. A linha vertical que corta o eixo das abcissas entre 10 e 11 unidades, representa a média geral do experimento, portanto o ambiente 3 (A3) é o mais próximo da média geral. Em relação aos genótipos, o

Figura 2.3. *Biplot* de um experimento com quatro genótipos em três ambientes a) $AMMI_1$; b) $AMMI_2$.

G3 apresenta maior resposta média, e o G1, a menor. Os genótipos 2 e 4 (G2 e G4) apresentam resposta média semelhante a média geral do experimento.

A primeira componente principal é aquela que retém a maior variabilidade dos dados. Se o percentual explicado pela CP_1 , isto é, a proporção da $SQ_{G \times E}$ representada no eixo, fosse suficientemente grande, o *biplot* $AMMI_1$ seria suficiente para caracterizar, sozinho, os genótipos e ambientes mais estáveis. Assim, no caso a) da Figura 2.3, os genótipos e ambientes próximos a linha horizontal que corta o eixo das ordenadas em zero (A3 e G3) seriam considerados os mais estáveis, isto porque escores baixos, ou seja, próximos de zero, são próprios de genótipos e ambientes que contribuíram pouco ou quase nada para a interação, sendo caracterizados como estáveis (Duarte e Vencovsky, 1999). Entretanto, geralmente é necessária mais de uma componente principal para fundamentar essa interpretação, assim o *biplot* $AMMI_1$ precisa ser interpretado simultaneamente com os demais eixos retidos na análise AMMI. Na Figura 2.3, em que foram consideradas apenas as duas primeiras componentes principais, se considera conjuntamente a interpretação do $AMMI_1$ e $AMMI_2$ para a identificação de genótipos e ambientes estáveis. No caso b) da figura, os genótipos e ambientes próximos a origem do gráfico, local em que as retas se cruzam, são considerados mais estáveis. Portanto, confirma-se a informação sobre a estabilidade do genótipo G3, enquanto que a caracterização do ambiente A3 como estável depende do percentual explicado pelas duas primeiras componentes principais, isto porque o ambiente A3 apresenta escores baixos (próximos de zero) para a primeira componente principal, porém, para a segunda CP, apresenta escores próximos ao valor um.

A estabilidade de um genótipo é um indicativo de sua adaptabilidade aos ambientes de teste, entretanto, na recomendação genotípica, a adaptabilidade é desejada concomitantemente com o seu desempenho médio. Supondo que valores maiores são desejáveis no exemplo da Figura 2.3, conclui-se que o G3 é um genótipo de interesse, pois apresenta alto desempenho médio e foi classificado como estável pela análise AMMI.

No entanto, os genótipos que não foram considerados estáveis, podem ser igualmente explorados na recomendação genotípica, isto porque, mesmo não apresentando ampla adaptabilidade, eles podem assumir adaptabilidades específicas a determinados ambientes, proporcionando uma recomendação regionalizada. Essas relações provenientes da interação $G \times E$ são identificadas no *biplot* observando-se os dois escores para cada par (genótipo, ambiente), de forma que genótipos e ambientes com escores de mesmo sinal são aqueles que interagem positivamente, evidenciando uma adaptabilidade específica, enquanto que genótipos e ambientes com sinais opostos são aqueles que interagem negativamente, sugerindo uma combinação desfavorável entre eles (Duarte e Vencovsky, 1999). Assim, retornando à Figura 2.3, caso b), constata-se que o genótipo G3 e ambiente A3, o genótipo G1 e o ambiente A1 e o genótipo G2 e o

ambiente A2 interagem positivamente, sugerindo uma adaptação específica entre eles.

2.7 Interação genótipos $\times$ ambientes para dados multiatributos

Os ensaios multiambientais são uma etapa fundamental do melhoramento de culturas, dado que, por meio da análise da interação genótipos $\times$ ambientes é possível identificar as cultivares superiores. Em geral, essa análise se baseia em uma única característica de interesse, como a comparação do rendimento de um conjunto de genótipos em vários ambientes. Entretanto, a análise dessa única característica não é suficiente para a recomendação de um cultivar.

Um experimento agrícola pode depender de muitas variáveis além do rendimento. No melhoramento de plantas, características físicas, químicas e sensoriais do cultivar também são importantes para a seleção e recomendação de um genótipo.

Exemplos de estudos da GEI com base em diferentes características biológicas de plantas podem ser vistos em López-Morales et al. (2017), em que foram avaliados, além do rendimento, características físicas como o índice de flotação, que é uma medida indireta da dureza dos grãos, e o peso hectolitro, que é influenciado por uniformidade, forma, densidade e tamanho do grão de milho em um ensaio com três genótipos em seis ambientes e em Ngailo et al. (2019), que analisaram a produtividade e a resistência à doença da batata-doce em 26 genótipos de batata-doce em seis ambientes. Ambos os estudos foram analisados por meio de análises AMMI individuais para cada característica.

Outros autores como Matkovic Stojšin et al. (2018), que analisaram o peso e o teor de proteína de dez genótipos de trigo em dois ambientes, e Koundinya et al. (2019), que analisaram informações sobre número de frutos por planta, peso do fruto e produção de frutos por planta, além de características químicas como o total de compostos fenólicos, em seu experimento com 40 cultivares de berinjela em quatro estações (ambientes), fizeram uso da análise AMMI individual para cada característica e a técnica estatística multivariada da análise de agrupamento.

A análise AMMI feita para cada característica individualmente não é suficiente para evidenciar toda a complexidade dos fatores envolvidos no estudo, uma vez que pode existir uma relação de sinergismo ou antagonismo entre diferentes caracteres biológicos, o que indica a necessidade de realizar uma análise simultânea das variáveis de interesse.

Nos últimos anos, algumas metodologias têm sido difundidas para os ensaios do melhoramento de plantas em tabelas de tripla entrada, em geral, para análise simultânea de genótipos, locais e anos/épocas. Desta forma os dados são dispostos em um arranjo tridimensional em que a observação x_{ijk} refere-se ao valor correspondente ao genótipo i ($i = 1, \dots, g$) no local j ($j = 1, \dots, l$) e no ano k ($k = 1, \dots, a$).

Nessa estrutura em que se considera a relação genótipo $\times$ local $\times$ ano, o pesquisador pode identificar, por exemplo, uma variabilidade constante dos locais em relação aos anos ou relações específicas entre genótipos e locais que diferem durante os anos do experimento.

Um procedimento comumente utilizado nesses casos é a transformação dessa estrutura tripla em uma matriz de dupla entrada, avaliando, por exemplo, a interação entre genótipos e locais para cada ano separadamente. É evidente que esse procedimento compromete a complexidade da informação, sendo incerta a interpretação global para todos os anos do estudo. Outro procedimento também comum nos ensaios em que ocorrem em múltiplos locais e épocas/anos é denominar cada combinação de locais e anos como um ambiente, como feito por Fan et al. (2007), que avaliaram treze híbridos de milho em dois anos e em dez locais, em que cada combinação de local-ano foi considerada um ambiente, ocasionando um total de 20 ambientes.

Silva et al. (2011) em seu estudo sobre a adaptabilidade e estabilidade de populações (genótipos) de cenoura, considerando dois anos e cinco locais, optaram por trabalhar com a média dos dois anos para

a metodologia AMMI, após constatarem que a interação genótipos $\times$ anos não foi significativa pela análise de variância conjunta.

Entretanto, para que não haja perda de informação, é necessário o uso metodologias que permitam o desdobramento e a interpretação da interação tripla. Varela et al. (2006) utilizaram a metodologia de componentes principais de três modos (*three mode principal components analysis* – 3MPCA), também denominada Tucker3, para analisar um banco de dados com 20 genótipos e quatro regimes de irrigação em três anos. O modelo Tucker3 também foi utilizado por Araújo et al. (2010), os quais compararam os resultados da análise AMMI obtida pela combinação de locais-épocas com o modelo Tucker3 em um ensaio com treze genótipos de feijão, concluindo que o modelo Tucker3 é eficiente e adequado para lidar com o arranjo triplo. Estes autores trabalharam com generalizações do modelo AMMI para arranjos de tripla entrada.

O uso da metodologia de componentes principais de três modos não se limita a relação genótipo $\times$ local $\times$ ano. Kroonenberg e Basford (1989) utilizaram a análise de componentes principais de três modos (3MPCA) para investigar a adaptabilidade de 58 variedades de soja medidas em quatro locais distintos durante dois anos. O experimento foi instalado em um delineamento em aleatorizado em blocos e foram considerados seis atributos diferentes na análise. Este método também foi utilizado por Araújo et al. (2010), os quais aplicaram a metodologia em um banco de dados proveniente de um experimento com 25 genótipos de trigo avaliados em quatro locais em que foram avaliados oito atributos, analisando assim a relação genótipo $\times$ local $\times$ atributo e ainda por Araújo et al. (2011) em um experimento com 36 genótipos de milho, 4 localidades e 10 atributos.

Basford e McLachlan (1985) propuseram o método de mistura de máxima verossimilhança de agrupamento (*mixture maximum likelihood method of clustering* - MIXCLUS). O MIXCLUS consiste em um método de agrupamento para tabelas de tripla entrada. Este método foi utilizado, conjuntamente com o 3MPCA, por Basford et al. (1990) para investigar a interação entre 25 variedades de algodão em nove ambientes e quatro atributos, e por Garcia-Peña et al. (2016) para estudar a GEI em 20 genótipos de girassol em oito ambientes, considerando três atributos diferentes. Segundo Basford et al. (1990), quando aplicados em conjunto, os métodos melhoram a utilidade um do outro na interpretação das informações sobre os padrões de resposta dos genótipos nos ambientes.

Mais recentemente, Garcia-Peña e Dias (2009) utilizaram a análise de Procrustes na comparação dos resultados obtidos pelo método AMMI para cada atributo em um conjunto de dados GEI com 19 genótipos de feijão avaliados em 18 ambientes considerando dois atributos.

Diante do exposto, constata-se a ausência de uma metodologia unificada para análise da interação GEI multiatributos. Alguns procedimentos adotados partem de análises AMMI individuais e recorrem a um método que permita a interpretação globalizada dos resultados e outros estudos trabalham com metodologias de três vias e generalizações do modelo AMMI para arranjos triplos, como as utilizadas por Varela et al. (2006) e Araújo et al. (2010).

2.8 Metodologias para dados de três modos

As metodologias aplicadas em dados classificados em três modos, como genótipo, ambiente e atributo ($G \times E \times A$) podem ser denominados métodos de três modos, três entradas ou, ainda, três vias, usando a terminologia de Carroll e Arabie (1980). De acordo com Kroonenberg (2008), desde o início da década de 1960 várias técnicas foram desenvolvidas especificamente para dados de três vias e essas técnicas foram posteriormente estendidas para dados de quatro vias ou mais. Entre essas técnicas estão a análise de componentes principais de três modos, análise de cluster de três vias e análise generalizada de Procrustes.

Em geral, as metodologias no contexto de dados de três vias podem ser distinguidas em duas classes: ordenação e técnicas de agrupamento. De acordo com Kruskal (1977) e Arabie et al. (1987), os dois tipos são em grande parte complementares, e fazem uso da mesma informação de maneiras diferentes. As análises indicadas por Kroonenberg (2008) para os dados de perfil são Tucker e Parafac (métodos de ordenação) e *Three-mode mixture method of clustering* e *three-mode partitioning* (métodos de agrupamento).

De acordo com Kroonenberg (2008), Lohmoller (1979) foi provavelmente um dos primeiros a discutir modelos aditivos e multiplicativos para dados de três vias. Ele sugeriu a seguinte generalização do modelo biaditivo:

$$\hat{x}_{ijk} = \mu + \beta_j + \gamma_k + \beta\gamma_{jk} + s_{jk} \times e_{ijk}$$

com a usual notação da análise de variância para a constante (geralmente a média geral) dada por μ , o efeito da variável β_j , o efeito da condição γ_k e o efeito combinado de interação variável-condição $\beta\gamma_{jk}$. O restante da variabilidade (s_{jk}) é decomposto por meio do modelo de componentes principais de três modos, ou Tucker3. Desde Lohmoller (1979), outros modelos aditivos e multiplicativos para dados de três vias foram propostos. Os dois principais modelos para decompor a variabilidade em um arranjo triplo são o modelo fatorial univariado de três modos e o modelo fatorial multivariado de dois modos.

O modelo fatorial univariado de três modos é o modelo univariado de análise de variância de efeitos principais para a parte aditiva e o modelo de Tucker3 para a parte multiplicativa, dado por

$$x_{ijk} = \mu + \alpha_i + \beta_j + \gamma_k + \alpha\beta_{ij} + \alpha\gamma_{ik} + \beta\gamma_{jk} + \alpha\beta\gamma_{ijk} \quad (2.13)$$

em que μ é a constante comum do experimento, α_i é o efeito principal do primeiro modo no ponto i , β_j é o efeito principal do segundo modo no ponto j e γ_k é o efeito principal do terceiro modo no ponto k . As interações duplas são dadas por $\alpha\beta_{ij}$, $\alpha\gamma_{ik}$ e $\beta\gamma_{jk}$ e a interação tripla $\alpha\beta\gamma_{ijk}$ é decomposta por Tucker3. Variações deste modelo ocorrem quando alguns dos termos aditivos são absorvidos pelo modelo de três modos ou quando as interações duplas são (parcialmente) decompostas por decomposições de valores singulares. Esse modelo foi utilizado para análise de melhoramento de plantas para dados de rendimento de genótipo por ano e por local.

O modelo fatorial multivariado de dois modos é dado por

$$x_{ik}^j = \mu^j + a_i^j + \gamma_k^j + \alpha\gamma_{ik}^j \quad (2.14)$$

em que o sobrescrito refere-se ao segundo modo. Este modelo é uma alternativa para dados em que um dos modos representa os atributos.

Para estimar os parâmetros do modelo fatorial multivariado de dois modos, podemos expressá-lo em termos de médias, isto é:

$$x_{ijk} = x_{ik}^j - \bar{x}_{.k}^j = (\bar{x}_{i.}^j - \bar{x}_{..}^j) + (\bar{x}_{ik}^j - \bar{x}_{i.}^j - \bar{x}_{.k}^j + \bar{x}_{..}^j) \quad (2.15)$$

em que o último termo no lado direito da equação (2.15), $(\bar{x}_{ik}^j - \bar{x}_{i.}^j - \bar{x}_{.k}^j + \bar{x}_{..}^j)$, descreve as diferenças relativas entre os indivíduos para o segundo modo. Se este termo for suficientemente pequeno, essa equação é aproximadamente igual a equação (2.12) e os resultados se assemelharão aos resultados de uma análise AMMI dos valores médios de todos os indivíduos. Portanto, é o tamanho desse termo que é crucial para a aplicação significativa de análises de três modos (Kroonenberg, 2008).

2.9 Análise de variância para três fatores

Considerando um estudo da GEI em que foram mensurados a atributos ($k = 1, 2, \dots, a$) em g genótipos ($i = 1, 2, \dots, g$) e e ambientes ($j = 1, 2, \dots, e$) com r repetições ($l = 1, 2, \dots, r$), o modelo clássico da análise de variância (ANOVA) pode ser dado por:

$$y_{ijkl} = \mu + g_i + e_j + a_k + (ge)_{ij} + (ga)_{ik} + (ea)_{jk} + (gea)_{ijk} + \varepsilon_{ijkl} \quad (2.16)$$

em que

y_{ijkl} é o valor observado do i -ésimo genótipo no j -ésimo ambiente, no k -ésimo atributo e na l -ésima repetição, com $i = 1, 2, \dots, g$, $j = 1, 2, \dots, e$, $k = 1, 2, \dots, a$ e $l = 1, 2, \dots, r$;

μ é uma constante, geralmente a média;

g_i é o efeito do i -ésimo genótipo;

e_j é o efeito do j -ésimo do ambiente;

a_k é o efeito do k -ésimo do atributo;

$(ge)_{ij}$ é o efeito da interação do i -ésimo genótipo com o j -ésimo ambiente;

$(ga)_{ik}$ é o efeito da interação do i -ésimo genótipo com o k -ésimo atributo;

$(ea)_{jk}$ é o efeito da interação do j -ésimo ambiente com o k -ésimo atributo;

$(gea)_{ijk}$ é o efeito da interação do i -ésimo genótipo com o j -ésimo ambiente com o k -ésimo atributo;

ε_{ijkl} é o erro experimental assumindo $\varepsilon_{ijkl} \sim N(0, \sigma^2)$ e $cov(\varepsilon_{ijkl}, \varepsilon_{i'j'k'l'}) = 0$ para $(i, j, k, l) \neq (i', j', k', l')$.

A Tabela 2.2 apresenta o esquema do quadro da análise de variância do experimento, assumindo um delineamento experimental inteiramente aleatorizado.

Tabela 2.2. Esquema da análise de variância para a atributos em g genótipos avaliados, e ambientes e r repetições.

FV	GL
G	$(g - 1)$
E	$(e - 1)$
A	$(a - 1)$
G×E	$(g - 1)(e - 1)$
G×A	$(g - 1)(a - 1)$
E×A	$(e - 1)(a - 1)$
G×E×A	$(g - 1)(e - 1)(a - 1)$
Resíduo	$gear - 1$
Total	$gear - 1$

Admitindo-se as condições marginais $\sum_{i=1}^g g_i = \sum_{j=1}^e e_j = \sum_{k=1}^a a_k = \sum_{i=1}^g \sum_{j=1}^e (ge)_{ij} = \sum_{i=1}^g \sum_{k=1}^a (ga)_{ik} = \sum_{j=1}^e \sum_{k=1}^a (ea)_{jk} = \sum_{i=1}^g \sum_{j=1}^e \sum_{k=1}^a (gea)_{ijk} = 0$, as hipóteses testadas no modelo (2.16) são:

- i. $H_0 : g_i = 0$, para todo i
 $H_1 : g_i \neq 0$, para algum i
- ii. $H_0 : e_j = 0$, para todo j
 $H_1 : e_j \neq 0$, para algum j

- iii. $H_0 : a_k = 0$, para todo k
 $H_1 : a_k \neq 0$, para algum k
- iv. $H_0 : (ge)_{ij} = 0$, para todo i, j
 $H_1 : (ge)_{ij} \neq 0$, para algum i, j
- v. $H_0 : (ga)_{ik} = 0$, para todo i, k
 $H_1 : (ga)_{ik} \neq 0$, para algum i, k
- vi. $H_0 : (ea)_{jk} = 0$, para todo j, k
 $H_1 : (ea)_{jk} \neq 0$, para algum j, k
- vii. $H_0 : (gea)_{ijk} = 0$, para todo i, j, k
 $H_1 : (gea)_{ijk} \neq 0$, para algum i, j, k

2.10 Decomposição da interação tripla

Para decompor a interação tripla é necessário a utilização de um método *multiway*. As análises *multiway* são aquelas adequadas a dados que envolvem vários fatores (ou vias, sendo também denominados dados multivias).

Os arranjos de dados de duas entradas, ou *two-way*, são os mais comuns e são representados por uma matriz que apresenta nas colunas as variáveis e nas linhas os objetos. Os arranjos de três entradas (*three-way*) são representados por um cubo, e, no lugar de linhas e colunas, consideram-se as fatias horizontais, verticais e frontais (Figura 2.4). De acordo com Smilde et al. (2004), existem três maneiras diferentes de fatiar um arranjo de três entradas $\mathbf{X}(I \times J \times K)$. A primeira maneira origina as fatias horizontais $\mathbf{X}_{i=1}, \mathbf{X}_{i=2}, \dots, \mathbf{X}_{i=I}$ com todas as dimensões $(J \times K)$, a segunda maneira origina as fatias verticais $\mathbf{X}_{j=1}, \mathbf{X}_{j=2}, \dots, \mathbf{X}_{j=J}$ com todas as dimensões $(I \times K)$ e a terceira maneira origina as fatias frontais $\mathbf{X}_{k=1}, \mathbf{X}_{k=2}, \dots, \mathbf{X}_{k=K}$ com todas as dimensões $(I \times J)$. O símbolo ' _ ' abaixo de $\mathbf{X}$, indica que é um arranjo de três entradas.

Figura 2.4. Particionamento de um arranjo de três entradas em fatias horizontais, verticais e frontais.

A decomposição de um arranjo de três entradas foi inicialmente proposta por Tucker (1964). Neste artigo, a decomposição de uma matriz de três vias nas matrizes de carga $\mathbf{A}$, $\mathbf{B}$ e $\mathbf{C}$ e um núcleo de três vias $\mathbf{G}$ é introduzida com um exemplo numérico hipotético de 12 indivíduos, 9 características e 5 avaliadores. A matemática do primeiro artigo de Tucker não estava completa e um novo artigo em 1966

(Tucker, 1966) foi publicado, no qual foi introduzido o produto de Kronecker. Os modelos apresentados por Tucker em seus artigos ficaram conhecidos como Componentes Principais de n Modos ou modelo de Tucker. Levin (1965) mostrou a semelhança entre o núcleo de uma decomposição de Tucker e os valores singulares em uma decomposição por valores singulares.

2.11 Modelo Tucker3

Uma generalização do modelo de componentes principais para dados de duas entradas considerando a decomposição por valores singulares de uma matriz $\mathbf{X}(I \times J)$ pode ser escrita como:

$$x_{ij} = \sum_{p=1}^P \sum_{q=1}^Q \tilde{a}_{ip} \tilde{g}_{pq} \tilde{b}_{jq} + e_{ij} \quad (2.17)$$

ou, matricialmente

$$\mathbf{X} = \mathbf{A} \tilde{\mathbf{G}} \mathbf{B}^T + \mathbf{E}$$

com $\tilde{\mathbf{G}}$ uma matriz núcleo não diagonal, em que ' $\sim$ ' acima de $\mathbf{G}$ é utilizado para indicar a diferença entre as matrizes núcleo convencionais e que $\tilde{g}_{pq}$ é um elemento da matriz $\tilde{\mathbf{G}}$. Este modelo não exige que $\mathbf{A}$ e $\mathbf{B}$ tenham as mesmas dimensões, diferentemente do que ocorre na decomposição por valores singulares, assim a matriz núcleo $\tilde{\mathbf{G}}$ é de dimensão $(P \times Q)$, de forma que o número de componentes pode ser diferente nos dois modos. Na análise tradicional de componentes principais, os vetores de cargas interagem aos pares, por exemplo, o primeiro vetor de escores interage com o primeiro vetor de cargas por uma magnitude definida como o primeiro valor singular. No modelo (2.17) todos os vetores podem interagir (Smilde et al., 2004).

A equação (2.17) pode ser generalizada para um arranjo triplo $\underline{\mathbf{X}}$ de dimensões $I \times J \times K$ com elementos x_{ijk} dados por

$$x_{ijk} = \sum_{p=1}^P \sum_{q=1}^Q \sum_{r=1}^R a_{ip} b_{jq} c_{kr} g_{pqr} + e_{ijk} \quad (2.18)$$

em que

x_{ijk} denota o elemento do arranjo $\underline{\mathbf{X}}$;

a_{ip} é um elemento da matriz de componentes principais $\mathbf{A}$ ($I \times P$) associada ao primeiro modo;

b_{jq} é um elemento da matriz de componentes principais $\mathbf{B}$ ($J \times Q$) associada ao segundo modo;

c_{kr} é um elemento da matriz de componentes principais $\mathbf{C}$ ($K \times R$) associada ao terceiro modo;

g_{pqr} é o elemento do arranjo núcleo $\underline{\mathbf{G}}$ ($P \times Q \times R$) que representa a relação entre o p -ésimo componente do primeiro modo com o q -ésimo componente do segundo modo e com o r -ésimo componente do terceiro modo;

e_{ijk} representa os elementos do arranjo de resíduos $\underline{\mathbf{E}}$ ($I \times J \times K$);

com $i = 1, \dots, I$, $j = 1, \dots, J$, $k = 1, \dots, K$, $p = 1, \dots, P$, $q = 1, \dots, Q$ e $r = 1, \dots, R$ (Timmerman e Kiers, 2000; Araújo et al., 2011).

O modelo dado em (2.18) é conhecido como Tucker3 de $\underline{\mathbf{X}}(P, Q, R)$, em que a notação (P, Q, R) é utilizada para indicar que o modelo tem P , Q e R fatores em três diferentes modos (ou entradas) (Smilde et al., 2004). A Figura 2.5 representa graficamente o modelo Tucker3.

Utilizando a notação do produto de Kronecker, o modelo de Tucker3 pode ser reescrito como (Smilde et al., 2004)

Figura 2.5. Representação gráfica do modelo Tucker3.

$$\mathbf{X} = \mathbf{A}\mathbf{G}(\mathbf{C} \otimes \mathbf{B})^T + \mathbf{E} \quad (2.19)$$

em que $\mathbf{G} = [\mathbf{G}_1 \mathbf{G}_2 \dots \mathbf{G}_R]$ é um desdobramento do arranjo núcleo $\underline{\mathbf{G}}$ com dimensão $(P \times QR)$, $\mathbf{G}_r$ é a r -ésima fatia frontal de dimensão $(P \times Q)$ de $\mathbf{G}$ e $\mathbf{X} = [\mathbf{X}_1 \mathbf{X}_2 \dots \mathbf{X}_k]$ é uma matriz $(I \times JK)$ sendo $\mathbf{X}_k$ a k -ésima fatia frontal $(I \times J)$ do arranjo de três entradas $\mathbf{X}(I \times J \times K)$.

2.11.1 Estimação do modelo Tucker3

Originalmente o algoritmo proposto por Tucker (1966) para ajustar o modelo não utilizava estimadores de mínimos quadrados e, portanto, apresentava propriedades pouco claras (Timmerman e Kiers, 2000; Varela e Torres, 2005). Para solucionar esse problema Kroonenberg e Leeuw (1980) introduziram uma metodologia alternativa para obtenção dos estimadores de mínimos quadrados com vetores de cargas ortogonais para o modelo. Esse algoritmo de estimação é denominado TUCKALS3.

Dado o arranjo $\underline{\mathbf{X}}$ de dimensão $I \times J \times K$, para ajustar o modelo de Tucker3 com dimensões $P \times Q \times R$ com matrizes de cargas ortogonais, deve-se, pelo algoritmo TUCKALS3, fornecer os valores iniciais da matriz $\mathbf{B}$ de dimensões $J \times Q$ e da matriz $\mathbf{C}$ de dimensões $K \times R$: os valores iniciais para a matriz $\mathbf{B}$ podem ser tomados como os Q primeiros vetores singulares-coluna de $\mathbf{X}_{(J \times IK)}$ e os valores iniciais para a matriz $\mathbf{C}$ serão os R primeiros vetores singulares-coluna de $\mathbf{X}_{(K \times IJ)}$. Os vetores singulares-coluna são obtidos da matriz $\mathbf{U}$ da decomposição por valores singulares (DVS) das matrizes de $\mathbf{B}$ e $\mathbf{C}$.

A matriz $\mathbf{A}$ é tomada como os P primeiros vetores singulares-coluna de $\mathbf{X}_{(I \times JK)}(\mathbf{C} \otimes \mathbf{B})$, portanto, $(\mathbf{C} \otimes \mathbf{B})$ tem JK linhas e RQ colunas. Como são considerados apenas os primeiros P vetores singulares-coluna de $\mathbf{X}_{(I \times JK)}(\mathbf{C} \otimes \mathbf{B})$, então a matriz $\mathbf{A}$ será de dimensão $(I \times P)$.

Atualizado a matriz $\mathbf{A}$, a matriz $\mathbf{B}$ é obtida tomando os Q primeiros vetores singulares-coluna de $\mathbf{X}_{(J \times IK)}(\mathbf{C} \otimes \mathbf{A})$ e, com os valores atualizados de $\mathbf{B}$, a matriz $\mathbf{C}$ será obtida tomando os R primeiros vetores singulares-coluna de $\mathbf{X}_{(K \times IJ)}(\mathbf{B} \otimes \mathbf{A})$.

Cada vez que as três matrizes $(\mathbf{A}, \mathbf{B}$ e $\mathbf{C})$ forem atualizadas, calcula-se $\mathbf{G} = \mathbf{A}^T \mathbf{X}(\mathbf{C} \otimes \mathbf{B})$ de dimensão $P \times QR$ e a mudança no ajuste é utilizada como uma medida de convergência, podendo ser facilmente calculada por meio da soma de quadrados dos elementos de $\mathbf{G}$, que representa a soma dos quadrados explicado pelo modelo provisório (Smilde et al., 2004).

2.11.2 Determinação da dimensionalidade de um modelo Tucker3

Timmerman e Kiers (2000) propuseram um procedimento para determinar a dimensionalidade de um modelo Tucker3. Esse procedimento consiste, inicialmente, em selecionar todos os possíveis modelos de Tucker3 que satisfazem as seguintes condições:

- i. $P \leq QR$
- ii. $Q \leq PR$
- iii. $R \leq PQ$

Em seguida, seleciona-se, entre os modelos com o mesmo número total de componentes $S = P + Q + R$, o modelo que tem a maior proporção da soma de quadrados ajustada (SQ_S), ou equivalentemente, a menor SQ residual. Para comparar os modelos com diferentes número totais de componentes (S), calcula-se $dif_S = SQ_S - SQ_{S-1}$, considerando apenas os dif_S que são consecutivamente (em ordem decrescente) maiores.

Um critério de corte para os dif_S também foi proposto por Timmerman e Kiers (2000), de forma que os dif_S devem ser maiores do que a proporção média da variabilidade explicada, calculada sobre todos os possíveis valores de S , isto é

$$\frac{\|\underline{Z}\|}{S_{min}} = \frac{SQ_{G \times E \times A}/b}{\min(I; JK) + \min(J; IK) + \min(K; IJ) - 3} \quad (2.20)$$

assim, todo dif_S que apresentar um valor abaixo do critério de corte não deve ser levado em conta na seleção do modelo. Timmerman e Kiers (2000) ainda definiram um valor de relevância $b_S = dif_S / dif_{S*}$, em que dif_{S*} é o maior valor depois de dif_S . Portanto, deve-se selecionar o modelo que tem o valor de b_S mais alto e esse procedimento é denominado *DifFit*.

2.12 Gráfico *joint plot*

Apesar do *joint plot* na análise de três modos ser similar ao *biplot* padrão, podendo utilizar todos os seus princípios interpretativos, o *joint plot* é a construção de um *biplot* dos componentes de dois modos, dado um componente do terceiro modo ou o modo de referência (Kroonenberg, 2008).

Sendo $\mathbf{G}$ ($P \times QR$) um desdobramento do arranjo núcleo $\underline{\mathbf{G}}$ e sendo $\mathbf{G}_r$ a r -ésima fatia frontal de dimensão $P \times Q$ de $\mathbf{G}$, para construir um *joint plot* após uma análise de Tucker3, deve-se decompor, via decomposição por valores singulares, a fatia do arranjo núcleo $\mathbf{G}_r$ ($P \times Q$), de forma que

$$\mathbf{G}_r = \mathbf{U}_r \mathbf{\Lambda}_r \mathbf{V}_r^T$$

e os vetores singulares $\mathbf{U}_r$ e $\mathbf{V}_r^T$ são combinados com as matrizes $\mathbf{A}$ e $\mathbf{B}$, respectivamente, e

$$\mathbf{A}_r^* = \left(\frac{I}{J}\right)^{1/4} \mathbf{A} \mathbf{U}_r \mathbf{\Lambda}_r^{1/2}$$

$$\mathbf{B}_r^* = \left(\frac{J}{I}\right)^{1/4} \mathbf{B} \mathbf{V}_r \mathbf{\Lambda}_r^{1/2}$$

obtendo a matriz $\mathbf{\Delta}_r = \mathbf{A} \mathbf{G}_r \mathbf{B}^T = \mathbf{A}_r^* \mathbf{B}_r^{*T}$ de dimensão $I \times J$ com $r = 1, 2, \dots, R$.

As colunas das matrizes de componentes ajustadas referem-se aos eixos do *joint plot*. Se a matriz $\mathbf{G}_r$ não é quadrada, seu posto será $M = \min(P, Q)$ e somente M *joint plots* poderão ser exibidos. Uma vez que $\mathbf{A}_r^* \mathbf{B}_r^{*T} = \mathbf{\Delta}_r$, cada elemento δ_{ij}^r é igual ao produto interno de $\mathbf{a}_r^* \mathbf{b}_r^{*T}$ e isso proporciona um aumento na associação entre a linha i da matriz de componentes $\mathbf{A}$ e a linha j da matriz de componentes $\mathbf{B}$, controlada pela r -ésima fatia do arranjo núcleo. Ao exibir simultaneamente os dois modos em um gráfico, inferências visuais podem ser feitas sobre seus relacionamentos. O espaçamento e a ordenação dos sujeitos em uma variável correspondem aos tamanhos de seus produtos internos e, portanto, à importância relativa dessa variável para os sujeitos (Kroonenberg, 2008).

Varela et al. (2006) fornecem a interpretação do *joint plot*: suponha um gráfico projetado sobre a componente principal r do terceiro modo, de forma que no *joint plot* aparecerão todos os níveis dos dois primeiros modos. A partir da matriz $\mathbf{C}$ (matriz de componentes principais do terceiro modo), selecionam-se os níveis do terceiro modo com mais peso no r -ésimo componente, sendo eles positivos ou negativos.

Admitindo que a matriz $\mathbf{C}$ tenha um valor positivo e alto associado ao k -ésimo nível do terceiro modo, então proximidades entre os níveis do primeiro e do segundo modo (por exemplo, o nível i para o primeiro modo e o nível j para o segundo modo) indicam que a interação tripla entre o i -ésimo nível do primeiro modo, o j -ésimo nível do segundo modo e o k -ésimo nível do terceiro modo é positiva.

Suponha, por exemplo, que a Figura 2.6 represente o *joint plot* de uma análise para interação tripla entre genótipos (primeiro modo), ambientes (segundo modo) e tempo de semeadura (terceiro modo), na qual a primeira componente da matriz $\mathbf{C}$ representa um contraste entre os tempos T1 (score positivo) e T2 (score negativo). Portanto, as interações entre $G1 \times A3$ e entre $G2 \times A1$ no tempo T1 são positivas.

Figura 2.6. *Joint plot* projetado dentro da primeira componente do terceiro modo considerando um experimento com três genótipos, três ambientes e dois tempos.

Por outro lado, se o nível i do primeiro modo estiver muito distante do nível j do segundo modo, isso indica que a interação tripla associada à combinação do i -ésimo do primeiro modo, j -ésimo nível do segundo modo e k -ésimo nível do terceiro modo tem um efeito negativo. No exemplo anterior, isso significa que, no tempo T1, o G1 possui uma interação negativa com o A1 e o G2 possui uma interação negativa com o A3.

Admitindo agora que a matriz $\mathbf{C}$ tem um valor negativo elevado associado ao nível k do terceiro modo, então proximidades entre os níveis do primeiro e do segundo modo (sendo o i -ésimo nível para o primeiro modo e o j -ésimo nível para o segundo modo) indicam que a interação tripla entre o i -ésimo nível do primeiro modo, o j -ésimo nível do segundo modo e o k -ésimo nível do terceiro modo é negativa. Por outro lado, se o i -ésimo nível do primeiro modo estiver muito distante do j -ésimo nível do segundo modo, isso indica que a interação tripla associada à combinação do i -ésimo, j -ésimo e k -ésimo níveis tem um efeito positivo.

Considerando o exemplo anterior, conclui-se pela Figura 2.6 que as interações entre $G1 \times A3$ e entre $G2 \times A1$ no tempo T2 são negativas. E as interações entre $G1 \times A1$ e entre $G2 \times A3$ no tempo T2 são positivas.

Em geral, os níveis de um modo específico localizado no centro do *joint plot* são considerados como tendo um desempenho médio (com relação à interação tripla) nos outros modos, sendo este o caso do G3 e A2 na Figura 2.6.

2.13 Análise de Procrustes

A análise de Procrustes é uma técnica analítica que permite a comparação de duas representações gráficas por meio de uma medida numérica. A estatística de procruste é obtida em três passos: translação, rotação e dilatação. Como indicado na Figura 2.7, a translação consiste em um deslocamento de todos os pontos por meio de uma distância constante no mesmo sentido; a rotação é o deslocamento fixo de todos os pontos por meio de um ângulo constante, mantendo a distância de cada ponto ao centróide; e a dilatação é o encolhimento ou expansão de todos os pontos através de uma constante em uma linha reta do ponto ao centróide da configuração (Garcia-Peña e Dias, 2009; Krzanowski, 2000).

Figura 2.7. Passos para obtenção da estatística de Procrustes. a) figura inicial; b) translação; c) rotação; d) dilatação.

Sejam $\mathbf{X}$ de dimensão $(n \times q)$ e $\mathbf{Y}$ $(n \times p)$, com $p \geq q$, matrizes que correspondem às coordenadas de representações gráficas (ou configurações) bidimensionais com n pontos provenientes de uma mesma observação. Assume-se que o r -ésimo ponto da primeira configuração corresponde ao r -ésimo ponto da segunda configuração.

O objetivo é comparar as duas configurações, admitindo que as linhas das matrizes correspondam às mesmas observações e que $p = q$. Nos casos em que $p > q$, $p - q$ colunas de zeros devem ser adicionadas ao final da matriz $\mathbf{X}$, assim, $\mathbf{X}$ e $\mathbf{Y}$ passam a ter dimensão $n \times p$. Segundo Krzanowski (2000), uma medida simples do grau de coincidência das duas configurações é a soma de quadrados das distâncias entre os pontos no espaço $\mathbf{Y}$ e os pontos correspondentes no espaço $\mathbf{X}$, que pode ser definida por:

$$M^2 = \sum_{r=1}^n (\mathbf{y}_r - \mathbf{x}_r)^T (\mathbf{y}_r - \mathbf{x}_r) \quad (2.21)$$

em que $\mathbf{X} = [\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n]^T$, $\mathbf{Y} = [\mathbf{y}_1, \mathbf{y}_2, \dots, \mathbf{y}_n]^T$ e $\mathbf{x}_r$ e $\mathbf{y}_r$ são os vetores de coordenadas do r -ésimo ponto nos dois espaços. A medida apresentada em (2.21) apresenta um cálculo direto da distância entre os pontos e é afetada pela posição, a orientação e a escala de uma configuração que são arbitrárias em relação à outra.

A comparação de duas configurações representa uma comparação das relações internas dos n pontos que não deve ser afetada pela posição, a orientação e a escala de uma configuração. As operações de translação, rotação e dilatação da análise de Procrustes não modificam essas relações internas. A

reflexão é outra operação que pode ser feita sem que as relações internas dos n pontos sejam modificadas, porém a reflexão em q dimensões de uma configuração corresponde a uma rotação em $q + 1$ dimensões.

Assim, para comparação de duas configurações, estatística M^2 de Procrustes é calculada após a translação, rotação e dilatação dos pontos, de forma que a estatística calculada corresponde a uma medida de "falta de ajuste" de uma configuração à outra. De acordo com Krzanowski (2000), deve-se fixar a configuração $\mathbf{Y}$ e igualar a configuração com coordenadas $\mathbf{X}$ a esta. A correspondência entre as configurações será feita pelas operações de translação, rotação e dilatação (sequencialmente) de forma a minimizar o M^2 . Matricialmente, isso implica que os pontos em $\mathbf{X}$ devem ser translacionados, rotacionados e dilatados para uma nova coordenada $\mathbf{x}_r^*$ dada por (Cox e Cox, 2000):

$$\mathbf{x}_r^* = \rho \mathbf{A}^T \mathbf{x}_r + \mathbf{b} \quad (2.22)$$

em que a matriz $\mathbf{A}$ é ortogonal e resulta na rotação e, possivelmente, reflexão de $\mathbf{x}_r$, o vetor $\mathbf{b}$ produz a translação e o valor de ρ , a dilatação. A nova soma de quadrados das distâncias entre os pontos, que é dada por:

$$M^2 = \sum_{r=1}^n (\mathbf{y}_r - \rho \mathbf{A}^T \mathbf{x}_r - \mathbf{b})^T (\mathbf{y}_r - \rho \mathbf{A}^T \mathbf{x}_r - \mathbf{b}) \quad (2.23)$$

As operações de translação, rotação e dilatação buscam minimizar (2.23) são apresentadas no Anexo A. Por meio delas, obtém-se a estatística de Procrustes, dada por:

$$M^2 = 1 - \frac{\{tr(\mathbf{X}^T \mathbf{Y} \mathbf{Y}^T \mathbf{X})^{1/2}\}^2}{tr(\mathbf{X}^T \mathbf{X}) tr(\mathbf{Y}^T \mathbf{Y})}$$

que indica maior similaridade entre as configurações, quanto menor for o seu valor.

3 MATERIAL E MÉTODOS

3.1 Material

Os dados utilizados neste trabalho são provenientes de experimentos de avaliação do escurecimento de grãos em feijoeiro pelos métodos de escurecimento natural do grão e escurecimento acelerado. Os experimentos foram instalados em delineamento experimental aleatorizados em três blocos no ano de 2016. Foram avaliados 19 genótipos de feijoeiros do tipo carioca, em seis ambientes no Estado de São Paulo: Campinas (A) e Tatuí (B) na época da “seca”, Votuporanga (C) e Ribeirão Preto (D) no inverno e Mococa (E) e Campinas (F), na época das “águas”. Os genótipos avaliados foram: BRS Perola, CHC 01-175-1, CNFC 11-948, CNFC 11-954, Gen 4-1F-19P, Gen 12-2F-67, Gen 20-4F-129, Gen 45-2F-293P, Gen 78-1A-59, Gen 86-12A-122, Gen 90-4A-160, Gen 104-1A-291, Gen 106-4A-317, Gen 106-6A-319, Gen 107-14A-336, Gen 125-10A-510, IAC Milenio, IAC Sintonia e LP 11-363 rotulados sequencialmente de G1 a G19.

Os genótipos foram avaliados quanto à cor do tegumento dos grãos em função do valor de luminosidade pelo método de escurecimento natural em condições de prateleira e pelo método de escurecimento acelerado em condições de câmara com luz ultravioleta e fluorescente. No método de escurecimento natural os grãos foram acondicionados em embalagens plásticas tipo *zip lock* e armazenados em condições de prateleira com temperatura e umidade ambiente. A mudança gradativa da coloração dos grãos foi determinada aos 0, 30, 60 e 90 dias. No método de escurecimento acelerado foram instaladas duas lâmpadas ultravioletas e duas lâmpadas fluorescentes, de forma alternada, na câmara. Os grãos foram colocados em placas de Petri abertas, de 90 milímetros de diâmetro, a 18 centímetros abaixo das lâmpadas, sendo avaliados com 24, 48, 72 e 96 horas de exposição. Neste estudo, cada método é considerado um atributo, sendo denominados AT1, AT2, ..., AT8.

Os valores de luminosidade (L^*), que representa a escala de luminosidade de 0 (preto) a 100 (branco) pelo sistema CIELAB, das amostras foram registrados utilizando-se colorímetro (modelo CR-410, Konica Minolta, Osaka, Japão) e foram expressos pela média de cinco aferições para cada amostra.

3.2 Métodos

A análise dos dados consistiu em duas etapas: o estudo da interação dupla entre genótipos $\times$ ambientes para cada atributo e a análise da interação tripla entre genótipos $\times$ ambientes $\times$ atributos.

Na primeira etapa, denominada análise de dados de dupla entrada, foi realizada a análise de variância para dois fatores, cujo modelo foi descrito na seção (3.2.1), e o modelo AMMI foi aplicado para cada um dos oito atributos, conforme apresentado na seção (3.2.2).

Na segunda etapa, denominada análise de dados de tripla entrada, foi realizada, inicialmente, uma análise de variância para três fatores para estudar a interação genótipos $\times$ ambientes $\times$ atributos, cujo modelo foi descrito na seção (3.2.3). O desdobramento da interação tripla foi obtido por meio do modelo de Tucker3, explicado na seção (3.2.4), reproduzindo, assim, o modelo AMMI para interação tripla da seção (3.2.5). A análise de Procrustes dada na seção (3.2.6) foi utilizada para comparação entre os resultados das análises de dados das duas etapas.

A segunda etapa foi realizada para dois casos diferentes: no primeiro momento considerando todos os oito atributos simultaneamente e, em um segundo momento, considerando os últimos tempos de cada um dos dois métodos de escurecimento avaliados no experimento.

3.2.1 Análise de Variância para dois fatores

Com o objetivo de identificar se existe a interação entre os 19 genótipos com os 6 ambientes do experimento, será inicialmente realizada uma análise de variância (ANOVA) para dois fatores, assumindo o modelo:

$$y_{ijr} = \mu + g_i + e_j + (ge)_{ij} + b_r(e_j) + \varepsilon_{ijr} \quad (3.1)$$

em que

y_{ijr} é o valor de luminosidade observado no i -ésimo genótipo, j -ésimo ambiente e r -ésimo bloco, com $i = 1, 2, \dots, 19$, $j = 1, 2, \dots, 6$ e $r = 1, 2, 3$;

μ é uma constante, geralmente a média;

g_i é o efeito do i -ésimo genótipo;

e_j é o efeito do j -ésimo do ambiente;

$(ge)_{ij}$ é o efeito da interação do i -ésimo genótipo com o j -ésimo ambiente;

$b_r(e_j)$ é o efeito do r -ésimo bloco dentro do j -ésimo ambiente;

ε_{ijr} é o erro experimental assumindo $\varepsilon_{ijr} \sim N(0, \sigma^2)$ e $cov(\varepsilon_{ijr}, \varepsilon_{i'j'r'}) = 0$ para $(i, j, r) \neq (i', j', r')$.

O modelo na sua forma matricial com $g = 19$, $e = 6$ e $r = 3$ é dado por

$$\mathbf{y} = \mathbf{X}\boldsymbol{\Theta} + \boldsymbol{\varepsilon}$$

$$\begin{bmatrix} y_{(1,1,1)} \\ y_{(1,1,2)} \\ y_{(1,1,3)} \\ y_{(1,2,1)} \\ y_{(1,2,2)} \\ y_{(1,2,3)} \\ \vdots \\ y_{(1,6,3)} \\ y_{(2,1,1)} \\ \vdots \\ y_{(2,6,3)} \\ \vdots \\ y_{(19,1,1)} \\ y_{(19,1,2)} \\ y_{(19,1,3)} \\ \vdots \\ y_{(19,6,3)} \end{bmatrix} = \begin{bmatrix} 1 & 1 & 0 & \cdots & 0 & 1 & 0 & \cdots & 0 & 1 & 0 & \cdots & 0 & 1 & 0 & \cdots & 0 \\ 1 & 1 & 0 & \cdots & 0 & 1 & 0 & \cdots & 0 & 1 & 0 & \cdots & 0 & 0 & 0 & \cdots & 0 \\ 1 & 1 & 0 & \cdots & 0 & 1 & 0 & \cdots & 0 & 1 & 0 & \cdots & 0 & 0 & 0 & \cdots & 0 \\ 1 & 1 & 0 & \cdots & 0 & 0 & 1 & \cdots & 0 & 0 & 1 & \cdots & 0 & 0 & 1 & \cdots & 0 \\ 1 & 1 & 0 & \cdots & 0 & 0 & 1 & \cdots & 0 & 0 & 1 & \cdots & 0 & 0 & 0 & \cdots & 0 \\ 1 & 1 & 0 & \cdots & 0 & 0 & 1 & \cdots & 0 & 0 & 1 & \cdots & 0 & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & 1 & 0 & \cdots & 0 & 0 & 0 & \cdots & 1 & 0 & 0 & \cdots & 0 & 0 & 0 & \cdots & 1 \\ 1 & 0 & 1 & \cdots & 0 & 1 & 0 & \cdots & 0 & 0 & 0 & \cdots & 0 & 1 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & 0 & 1 & \cdots & 0 & 0 & 0 & \cdots & 1 & 0 & 0 & \cdots & 0 & 0 & 0 & \cdots & 1 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & 0 & 0 & \cdots & 1 & 1 & 0 & \cdots & 0 & 0 & 0 & \cdots & 0 & 1 & 0 & \cdots & 0 \\ 1 & 0 & 0 & \cdots & 1 & 1 & 0 & \cdots & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & \cdots & 0 \\ 1 & 0 & 0 & \cdots & 1 & 1 & 0 & \cdots & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & 0 & 0 & \cdots & 1 & 0 & 0 & \cdots & 1 & 0 & 0 & \cdots & 1 & 0 & 0 & \cdots & 1 \end{bmatrix} \begin{bmatrix} \mu \\ g_1 \\ g_2 \\ \vdots \\ g_{19} \\ e_1 \\ e_2 \\ \vdots \\ e_6 \\ (ge)_{(1,1)} \\ (ge)_{(1,2)} \\ \vdots \\ (ge)_{(19,6)} \\ b_1(e_1) \\ b_1(e_2) \\ \vdots \\ b_3(e_6) \end{bmatrix} + \begin{bmatrix} \varepsilon_{(1,1,1)} \\ \varepsilon_{(1,1,2)} \\ \varepsilon_{(1,1,3)} \\ \varepsilon_{(1,2,1)} \\ \varepsilon_{(1,2,2)} \\ \varepsilon_{(1,2,3)} \\ \vdots \\ \varepsilon_{(1,6,3)} \\ \varepsilon_{(2,1,1)} \\ \vdots \\ \varepsilon_{(2,6,3)} \\ \vdots \\ \varepsilon_{(19,1,1)} \\ \varepsilon_{(19,1,2)} \\ \varepsilon_{(19,1,3)} \\ \vdots \\ \varepsilon_{(19,6,3)} \end{bmatrix}$$

com $\mathbf{y}$ de ordem $(ger \times 1)$, $\mathbf{X}$ $(ger \times p)$, $\boldsymbol{\Theta}$ $(p \times 1)$ e $\boldsymbol{\varepsilon}$ $(ger \times 1)$ em que p é o número de parâmetros do modelo, isto é, $p = 1 + g + e + ge + re$.

A matriz $\mathbf{X}$ é composta pelas submatrizes $\mathbf{X}_T$, $\mathbf{X}_G$, $\mathbf{X}_E$, $\mathbf{X}_{GE}$ e $\mathbf{X}_B$, isto é

$$\mathbf{X}_T^T = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & \cdots & 1 & 1 & \cdots & 1 & \cdots & 1 & 1 & 1 & \cdots & 1 \end{bmatrix}$$

$$\begin{aligned}
\mathbf{X}_G^T &= \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & \cdots & 1 & 0 & \cdots & 0 & \cdots & 0 & 0 & 0 & \cdots & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & \cdots & 0 & 1 & \cdots & 1 & \cdots & 0 & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & 0 & 0 & 0 & \cdots & 0 & 0 & \cdots & 0 & \cdots & 1 & 1 & 1 & \cdots & 1 \end{bmatrix} \\
\mathbf{X}_E^T &= \begin{bmatrix} 1 & 1 & 1 & 0 & 0 & 0 & \cdots & 0 & 1 & \cdots & 0 & \cdots & 1 & 1 & 1 & \cdots & 0 \\ 0 & 0 & 0 & 1 & 1 & 1 & \cdots & 0 & 0 & \cdots & 0 & \cdots & 0 & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & 0 & 0 & 0 & \cdots & 1 & 0 & \cdots & 1 & \cdots & 0 & 0 & 0 & \cdots & 1 \end{bmatrix} \\
\mathbf{X}_{GE}^T &= \begin{bmatrix} 1 & 1 & 1 & 0 & 0 & 0 & \cdots & 0 & 0 & \cdots & 0 & \cdots & 0 & 0 & 0 & \cdots & 0 \\ 0 & 0 & 0 & 1 & 1 & 1 & \cdots & 0 & 0 & \cdots & 0 & \cdots & 0 & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & 0 & 0 & 0 & \cdots & 0 & 0 & \cdots & 0 & \cdots & 0 & 0 & 0 & \cdots & 1 \end{bmatrix} \\
\mathbf{X}_B^T &= \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & \cdots & 0 & 1 & \cdots & 0 & \cdots & 1 & 0 & 0 & \cdots & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & \cdots & 0 & 0 & \cdots & 0 & \cdots & 0 & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & 0 & 0 & 0 & \cdots & 1 & 0 & \cdots & 1 & \cdots & 0 & 0 & 0 & \cdots & 1 \end{bmatrix}
\end{aligned}$$

Os números de graus de liberdade e somas de quadrados associados aos efeitos do modelo, assumindo um delineamento aleatorizado em blocos, serão obtidos por meio dos Diagramas de Hasse das Figuras 3.1 e 3.2.

Figura 3.1. Diagrama de Hasse para os números de graus de liberdade associados aos termos do modelo (3.1)

Figura 3.2. Diagrama de Hasse para as somas de quadrados associadas aos termos do modelo (3.1)

A Tabela 3.1 apresenta o quadro da análise de variância do experimento. Os números de graus de liberdade foram obtidos pela diferença entre o número de graus de liberdade total e a soma dos números de graus de liberdade do bloco, genótipos, ambientes e interação. A mesma diferença foi utilizada para se obter a soma de quadrados do resíduo.

Tabela 3.1. Esquema da análise de variância para g genótipos avaliados em e ambientes com r blocos.

Fonte de variação	GL	SQ
Bloco d. Amb	$e(r - 1)$	$\mathbf{y}^T(\mathbf{M}_B - \mathbf{M}_E)\mathbf{y}$
Genótipos (G)	$(g - 1)$	$\mathbf{y}^T(\mathbf{M}_G - \mathbf{M}_T)\mathbf{y}$
Ambientes (E)	$(e - 1)$	$\mathbf{y}^T(\mathbf{M}_E - \mathbf{M}_T)\mathbf{y}$
Interação (G×E)	$(g - 1)(e - 1)$	$\mathbf{y}^T(\mathbf{M}_{GE} - \mathbf{M}_G - \mathbf{M}_E + \mathbf{M}_T)\mathbf{y}$
Resíduo	$e(g - 1)(r - 1)$	$\mathbf{y}^T(\mathbf{M}_U - \mathbf{M}_{GE} - \mathbf{M}_B + \mathbf{M}_E)\mathbf{y}$
Total	$ger - 1$	$\mathbf{y}^T(\mathbf{M}_U - \mathbf{M}_T)\mathbf{y}$

A matriz $\mathbf{M}_U$ é uma matriz identidade de dimensão $ger \times ger$ e as matrizes núcleo das somas de quadrados são dadas por

$$\begin{aligned}
 \mathbf{M}_T &= \mathbf{X}_T(\mathbf{X}_T^T \mathbf{X}_T)^{-1} \mathbf{X}_T^T \\
 \mathbf{M}_G &= \mathbf{X}_G(\mathbf{X}_G^T \mathbf{X}_G)^{-1} \mathbf{X}_G^T \\
 \mathbf{M}_E &= \mathbf{X}_E(\mathbf{X}_E^T \mathbf{X}_E)^{-1} \mathbf{X}_E^T \\
 \mathbf{M}_{GE} &= \mathbf{X}_{GE}(\mathbf{X}_{GE}^T \mathbf{X}_{GE})^{-1} \mathbf{X}_{GE}^T \\
 \mathbf{M}_B &= \mathbf{X}_B(\mathbf{X}_B^T \mathbf{X}_B)^{-1} \mathbf{X}_B^T
 \end{aligned}$$

Os quadrados médios para genótipos, ambientes, interação genótipos $\times$ ambientes, blocos dentro de ambientes e resíduo são dados pela razão entre as somas de quadrados e seus respectivos números de graus de liberdade, isto é

$$QM_G = \frac{SQ_G}{gl_G} = \frac{\mathbf{y}^T(\mathbf{M}_G - \mathbf{M}_T)\mathbf{y}}{g - 1}$$

$$\begin{aligned}
QM_E &= \frac{SQ_E}{gl_E} = \frac{\mathbf{y}^T(\mathbf{M}_E - \mathbf{M}_T)\mathbf{y}}{e - 1} \\
QM_{GE} &= \frac{SQ_{GE}}{gl_{GE}} = \frac{\mathbf{y}^T(\mathbf{M}_{GE} - \mathbf{M}_G - \mathbf{M}_E + \mathbf{M}_T)\mathbf{y}}{(g - 1)(e - 1)} \\
QM_B &= \frac{SQ_B}{gl_B} = \frac{\mathbf{y}^T(\mathbf{M}_B - \mathbf{M}_E)\mathbf{y}}{e(b - 1)} \\
QM_{Res} &= \frac{SQ_{Res}}{gl_{Res}} = \frac{\mathbf{y}^T(\mathbf{M}_U - \mathbf{M}_{GE} - \mathbf{M}_B + \mathbf{M}_E)\mathbf{y}}{e(g - 1)(r - 1)}
\end{aligned}$$

3.2.2 Modelo AMMI para um único atributo

Quando a interação genótipos $\times$ ambientes for significativa, segue-se a análise do modelo AMMI, aplicando a decomposição da $SQ_{G \times E}$. Em um experimento no delineamento aleatorizado em blocos, a observação y_{ijk} no modelo AMMI é dada pelas médias das respostas das combinações genótipos $\times$ ambientes feitas nos b blocos, isto é, $(\sum_k y_{ijk})/b$ (Gauch, 2013). No presente estudo, antes de aplicar a decomposição, é necessária a criação de uma matriz de ordem $g \times e$, isto é, 19×6 com as médias dos $b = 3$ blocos para cada uma das combinações de genótipos e ambientes.

A resposta média sobre b blocos do i -ésimo genótipo e j -ésimo ambiente pode ser modelada como:

$$y_{ij} = \mu + g_i + e_j + \sum_{k=1}^n \lambda_k \alpha_{ik} \gamma_{jk} + \rho_{ij} + \varepsilon_{ij} \quad (3.2)$$

em que

y_{ij} é a resposta média do i -ésimo genótipo no j -ésimo ambiente com $i = 1, 2, \dots, 19$ e $j = 1, 2, \dots, 6$;

μ é uma constante comum às respostas;

g_i e e_j correspondem, respectivamente, ao efeito principal do genótipo i e do ambiente j ;

λ_k é a raiz quadrada do k -ésimo autovalor das matrizes $(\mathbf{GE})(\mathbf{GE})^T$ e $(\mathbf{GE})^T(\mathbf{GE})$, com $(\mathbf{GE}) = [(\hat{g}e)_{ij}]$ dada em (2.10) e $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_p$ e $n < p$ em que n é o número de eixos ou componentes principais da interação $G \times E$;

α_{ik} é o i -ésimo valor estimado do vetor de escores genotípicos $\boldsymbol{\alpha}_k$ associado a λ_k ;

γ_{jk} é o j -ésimo valor estimado do vetor de escores ambientais $\boldsymbol{\gamma}_k$ associado a λ_k ;

ε_{ij} é o erro experimental associado à observação y_{ij} , assumindo $\varepsilon_{ij} \sim N(0, \sigma^2/b)$ e $cov(\varepsilon_{ij}, \varepsilon_{i'j'}) = 0$ para $(i, j) \neq (i', j')$.

O número de graus de liberdade para o modelo AMMI será determinado pelo método de Gollob, calculado como $GL_{CP_k} = g + e - 1 - 2k$, com $k = 1, 2, \dots, p$, $p = \min[(g-1), (e-1)]$ e GL_{CP_k} sendo o k -ésimo eixo das componentes principais.

A significância das componentes principais será testada por meio do quadrado médio de resíduo do modelo (3.1), para o qual a soma de quadrado da componente principal k será dada por $b\lambda_k^2$ em que b é o número de blocos (Onofri e Ciricifolo, 2007).

3.2.3 Análise de variância para três fatores

Considerando que foram mensurados 8 atributos em 19 genótipos e 6 ambientes em 3 blocos, o modelo assumido considerando a interação tripla, será dado por:

$$y_{ijk r} = \mu + g_i + e_j + a_k + b_r(e_j) + (ge)_{ij} + (ga)_{ik} + (ea)_{jk} + (gea)_{ijk} + \varepsilon_{ijk r} \quad (3.3)$$

em que

$y_{ijk r}$ é o valor observado do i -ésimo genótipo no j -ésimo ambiente, no k -ésimo atributo e no r -ésimo bloco, com $i = 1, 2, \dots, 19$, $j = 1, 2, \dots, 6$, $a = 1, 2, \dots, 8$ e $r = 1, 2, 3$;

μ é uma constante, geralmente a média;

g_i é o efeito do i -ésimo genótipo;

e_j é o efeito do j -ésimo do ambiente;

$b_r(e_j)$ é o efeito do r -ésimo bloco dentro do j -ésimo ambiente;

$(ge)_{ij}$ é o efeito da interação do i -ésimo genótipo com o j -ésimo ambiente;

$(ga)_{ik}$ é o efeito da interação do i -ésimo genótipo com o k -ésimo atributo;

$(ea)_{jk}$ é o efeito da interação do j -ésimo ambiente com o k -ésimo atributo;

$(gea)_{ijk}$ é o efeito da interação do i -ésimo genótipo com o j -ésimo ambiente com o k -ésimo atributo;

$\varepsilon_{ijk r}$ é o erro experimental assumindo $\varepsilon_{ijk r} \sim N(0, \sigma^2)$ e $cov(\varepsilon_{ijk r}, \varepsilon_{i'j'r'k'}) = 0$ para $(i, j, r, k) \neq (i', j', r', k')$.

O modelo matricial com $g = 19$, $e = 6$, $a = 8$ e $b = 3$ é dado por $\mathbf{y} = \mathbf{X}\boldsymbol{\Theta} + \boldsymbol{\varepsilon}$. A matriz $\mathbf{X}$ tem dimensão $(geab \times p)$ e é composta pelas submatrizes $\mathbf{X}_T$, $\mathbf{X}_G$, $\mathbf{X}_E$, $\mathbf{X}_A$, $\mathbf{X}_{GE}$, $\mathbf{X}_{GA}$, $\mathbf{X}_{EA}$, $\mathbf{X}_{GEA}$ e $\mathbf{X}_B$.

$$\mathbf{X}_T^T = \begin{bmatrix} 1 & 1 & 1 & 1 & \dots & 1 & 1 & \dots & 1 & \dots & 1 \end{bmatrix}$$

$$\mathbf{X}_G^T = \begin{bmatrix} 1 & 1 & 1 & 1 & \dots & 1 & 1 & \dots & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & 0 & \dots & 0 & 0 & \dots & 0 & \dots & 1 \end{bmatrix}$$

$$\mathbf{X}_E^T = \begin{bmatrix} 1 & 1 & 1 & 1 & \dots & 1 & 0 & \dots & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & 0 & \dots & 0 & 0 & \dots & 1 & \dots & 1 \end{bmatrix}$$

$$\mathbf{X}_A^T = \begin{bmatrix} 1 & 1 & 1 & 0 & \dots & 0 & 1 & \dots & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & 0 & \dots & 1 & 0 & \dots & 1 & \dots & 1 \end{bmatrix}$$

$$\mathbf{X}_{GE}^T = \begin{bmatrix} 1 & 1 & 1 & 1 & \dots & 1 & 0 & \dots & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & 0 & \dots & 0 & 0 & \dots & 0 & \dots & 1 \end{bmatrix}$$

$$\mathbf{X}_{GA}^T = \begin{bmatrix} 1 & 1 & 1 & 0 & \dots & 0 & 1 & \dots & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & 0 & \dots & 0 & 0 & \dots & 0 & \dots & 1 \end{bmatrix}$$

$$\mathbf{X}_{EA}^T = \begin{bmatrix} 1 & 1 & 1 & 0 & \dots & 0 & 0 & \dots & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & 0 & \dots & 0 & 0 & \dots & 1 & \dots & 1 \end{bmatrix}$$

$$\mathbf{X}_{GEA}^T = \begin{bmatrix} 1 & 1 & 1 & 0 & \cdots & 0 & 0 & \cdots & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & 0 & \cdots & 0 & 0 & \cdots & 0 & \cdots & 1 \end{bmatrix}$$

$$\mathbf{X}_B^T = \begin{bmatrix} 1 & 0 & 0 & 1 & \cdots & 0 & 0 & \cdots & 0 & \cdots & 0 \\ 0 & 0 & 0 & 0 & \cdots & 0 & 1 & \cdots & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & 0 & \cdots & 0 & 0 & \cdots & 1 & \cdots & 1 \end{bmatrix}$$

$$\mathbf{y} = \begin{bmatrix} y_{(1,1,1,1)} \\ y_{(1,1,1,2)} \\ y_{(1,1,1,3)} \\ y_{(1,1,2,1)} \\ \vdots \\ y_{(1,1,8,3)} \\ y_{(1,2,1,1)} \\ \vdots \\ y_{(1,6,8,3)} \\ \vdots \\ y_{(19,6,8,3)} \end{bmatrix} \quad \mathbf{\Theta} = \begin{bmatrix} \mu \\ g_1 \\ \vdots \\ g_{19} \\ e_1 \\ \vdots \\ e_6 \\ a_1 \\ \vdots \\ a_8 \\ (ge)_{(1,1)} \\ \vdots \\ (ge)_{(19,6)} \\ (ga)_{(1,1)} \\ \vdots \\ (ga)_{(19,8)} \\ (ea)_{(1,1)} \\ \vdots \\ (ea)_{(6,8)} \\ (gea)_{(1,1,1)} \\ \vdots \\ (gea)_{(19,6,8)} \\ b_1(e_1) \\ b_1(e_2) \\ \vdots \\ b_3(e_6) \end{bmatrix} \quad \mathbf{\varepsilon} = \begin{bmatrix} \varepsilon_{(1,1,1,1)} \\ \varepsilon_{(1,1,1,2)} \\ \varepsilon_{(1,1,1,3)} \\ \varepsilon_{(1,1,2,1)} \\ \vdots \\ \varepsilon_{(1,1,8,3)} \\ \varepsilon_{(1,2,1,1)} \\ \vdots \\ \varepsilon_{(1,6,8,3)} \\ \vdots \\ \varepsilon_{(19,6,8,3)} \end{bmatrix}$$

com $\mathbf{y}$ de ordem $(geab \times 1)$, $\boldsymbol{\Theta}$ $(p \times 1)$ e $\boldsymbol{\varepsilon}$ $(geab \times 1)$ em que p é o número de parâmetros do modelo, isto é, $p = 1 + g + e + a + b + ge + ga + ea + gea$.

Os números de graus de liberdade do experimento e as somas de quadrados associados aos termos do modelo (3.3) serão obtidos por meio dos Diagramas de Hasse das Figuras 3.3 e 3.4.

Figura 3.3. Diagrama de Hasse para os números de graus de liberdade associados aos termos do modelo (3.3)

Figura 3.4. Diagrama de Hasse para as somas de quadrados associadas aos termos do modelo (3.3)

Tabela 3.2. Esquema da análise de variância para experimentos com a atributos em g genótipos avaliados e e ambientes com b blocos.

FV	GL	SQ
B[A]	$e(b-1)$	$\mathbf{y}^T(M_B - M_E)\mathbf{y}$
G	$(g-1)$	$\mathbf{y}^T(M_G - M_T)\mathbf{y}$
E	$(e-1)$	$\mathbf{y}^T(M_E - M_T)\mathbf{y}$
A	$(a-1)$	$\mathbf{y}^T(M_A - M_T)\mathbf{y}$
G×E	$(g-1)(e-1)$	$\mathbf{y}^T(M_{GE} - M_G - M_E + M_T)\mathbf{y}$
G×A	$(g-1)(a-1)$	$\mathbf{y}^T(M_{GA} - M_G - M_A + M_T)\mathbf{y}$
E×A	$(e-1)(a-1)$	$\mathbf{y}^T(M_{EA} - M_E - M_A + M_T)\mathbf{y}$
G×E×A	$(g-1)(e-1)(a-1)$	$\mathbf{y}^T(M_{GEA} - M_{GE} - M_{GA} - M_{EA} + M_G + M_E + M_A - M_T)\mathbf{y}$
Resíduo	$e(ga-1)(b-1)$	$\mathbf{y}^T(M_U - M_{GEA} - M_B + M_E)\mathbf{y}$
Total	$geab-1$	$\mathbf{y}^T(M_U - M_T)\mathbf{y}$

A Tabela 3.2 apresenta o esquema do quadro da análise de variância do experimento. Os números de graus de liberdade e a soma de quadrados do resíduo foram obtidos por diferença.

A matriz $\mathbf{M}_U$ é uma matriz identidade de dimensão $geab \times geab$ e as matrizes núcleo das somas de quadrados serão dadas por

$$\begin{aligned}
\mathbf{M}_T &= \mathbf{X}_T(\mathbf{X}_T^T \mathbf{X}_T)^{-1} \mathbf{X}_T^T \\
\mathbf{M}_G &= \mathbf{X}_G(\mathbf{X}_G^T \mathbf{X}_G)^{-1} \mathbf{X}_G^T \\
\mathbf{M}_E &= \mathbf{X}_E(\mathbf{X}_E^T \mathbf{X}_E)^{-1} \mathbf{X}_E^T \\
\mathbf{M}_A &= \mathbf{X}_A(\mathbf{X}_A^T \mathbf{X}_A)^{-1} \mathbf{X}_A^T \\
\mathbf{M}_{GE} &= \mathbf{X}_{GE}(\mathbf{X}_{GE}^T \mathbf{X}_{GE})^{-1} \mathbf{X}_{GE}^T \\
\mathbf{M}_{GA} &= \mathbf{X}_{GA}(\mathbf{X}_{GA}^T \mathbf{X}_{GA})^{-1} \mathbf{X}_{GA}^T \\
\mathbf{M}_{EA} &= \mathbf{X}_{EA}(\mathbf{X}_{EA}^T \mathbf{X}_{EA})^{-1} \mathbf{X}_{EA}^T \\
\mathbf{M}_{GEA} &= \mathbf{X}_{GEA}(\mathbf{X}_{GEA}^T \mathbf{X}_{GEA})^{-1} \mathbf{X}_{GEA}^T \\
\mathbf{M}_B &= \mathbf{X}_B(\mathbf{X}_B^T \mathbf{X}_B)^{-1} \mathbf{X}_B^T
\end{aligned}$$

Os quadrados médios para as fontes de variação da Tabela 3.2 são dados pela razão entre a soma de quadrado e seus respectivos números de graus de liberdade.

3.2.4 Modelo Tucker3

Um algoritmo de mínimos quadrados para ajustar o modelo de Tucker3 com dimensões $P \times Q \times R$ com matrizes de cargas ortogonais para um arranjo $\underline{\mathbf{X}}$ ($I \times J \times K$) pode ser implementado como (Smilde et al., 2004):

1. Inicialize $\mathbf{B}$ e $\mathbf{C}$;
2. $\mathbf{A}$ igual a P vetores singulares-coluna de $\mathbf{X}_{(I \times JK)}(\mathbf{C} \otimes \mathbf{B})$;
3. $\mathbf{B}$ igual a Q vetores singulares-coluna de $\mathbf{X}_{(J \times IK)}(\mathbf{C} \otimes \mathbf{A})$;
4. $\mathbf{C}$ igual a R vetores singulares-coluna de $\mathbf{X}_{(K \times IJ)}(\mathbf{B} \otimes \mathbf{A})$;
5. Retorne ao passo 2. até que as mudanças no ajuste do modelo sejam relativamente pequenas;
6. Faça $\mathbf{G} = \mathbf{A}^T \mathbf{X}(\mathbf{C} \otimes \mathbf{B})$.

Para inicializar o algoritmo, deve-se fornecer valores iniciais de $\mathbf{B}$ e $\mathbf{C}$. Geralmente, eles são tomados como vetores singulares-coluna de uma decomposição por valores singulares de uma matriz, onde o modo fornecido é mantido como modo de linha e os dois modos adicionais são combinados no modo de coluna. Ou seja, os valores iniciais para a matriz $\mathbf{B}$ foram tomados como os Q primeiros vetores singulares-coluna de $\mathbf{X}_{(J \times IK)}$ e os valores iniciais para a matriz $\mathbf{C}$ foram tomados como os R primeiros vetores singulares-coluna de $\mathbf{X}_{(K \times IJ)}$.

Seja o modelo Tucker3 como dado em (2.19) no qual há interações entre todos os vetores de componentes em todos os modos. Cada elemento g_{pqr} do arranjo $\underline{\mathbf{G}}(P \times Q \times R)$ dá a importância da combinação de cada componente. Assim, a importância (ou variabilidade explicada) da componente p da matriz $\mathbf{A}$ será calculada por:

$$\frac{\sum_{q=1}^Q \sum_{r=1}^R g_{pqr}^2}{SQ_{G \times E \times A}/b} \times 100 \quad (3.4)$$

em que b é o número de blocos. De maneira semelhante, a importância da componente q da matriz $\mathbf{B}$ será calculada por:

$$\frac{\sum_{p=1}^P \sum_{r=1}^R g_{pqr}^2}{SQ_{G \times E \times A}/b} \times 100 \quad (3.5)$$

e a importância da componente r da matriz $\mathbf{C}$ será calculada por:

$$\frac{\sum_{p=1}^P \sum_{q=1}^Q g_{pqr}^2}{SQ_{G \times E \times A}/b} \times 100 \quad (3.6)$$

A importância da combinação da componente p da matriz $\mathbf{A}$, com a componente q da matriz $\mathbf{B}$ e a componente r da matriz $\mathbf{C}$ será calculada por:

$$\frac{g_{pqr}^2}{SQ_{G \times E \times A}/b} \times 100 \quad (3.7)$$

3.2.5 Modelo AMMI para a interação tripla

Em um experimento no delineamento aleatorizado em blocos, a resposta média sobre b blocos do i -ésimo genótipo, j -ésimo ambiente e k -ésimo atributo será

$$y_{ijk} = \mu + g_i + e_j + a_k + (ge)_{ij} + (ga)_{ik} + (ea)_{jk} + \sum_{p=1}^P \sum_{q=1}^Q \sum_{r=1}^R a_{ip} b_{jq} c_{kr} g_{pqr} + \rho_{ijk} + \varepsilon_{ijk} \quad (3.8)$$

com y_{ijk} , μ , g_i , e_j e a_k , $(ge)_{ij}$, $(ga)_{ik}$ e $(ea)_{jk}$ como definidos em (3.3);

a_{ip} é o i -ésimo elemento da p -ésima coluna da matriz de componentes $\mathbf{A}$;

b_{jq} é o j -ésimo elemento da q -ésima coluna da matriz de componentes $\mathbf{B}$;

c_{kr} é o k -ésimo elemento da r -ésima coluna da matriz de componentes $\mathbf{C}$;

ρ_{ijk} é o resíduo adicional da interação tripla que não é explicado pelo modelo Tucker3 (P,Q,R), sendo este modelo responsável pelo desdobramento da interação $(gea)_{ijk} = \sum_{p=1}^P \sum_{q=1}^Q \sum_{r=1}^R a_{ip} b_{jq} c_{kr} g_{pqr} + \rho_{ijk}$; ε_{ijk} o erro experimental associado à observação y_{ijk} , assumindo $\varepsilon_{ijk} \sim N(0, \sigma^2/b)$ e $cov(\varepsilon_{ijk}, \varepsilon_{i'j'k'}) = 0$ para $(i, j, k) \neq (i', j', k')$.

No arranjo de interações $\underline{\mathbf{Z}} = (g\hat{e}a)_{ijk}$ cada elemento $(g\hat{e}a)_{ijk}$ foi obtido por meio da relação

$$(g\hat{e}a)_{ijk} = Y_{ijk} - \bar{Y}_{ij.} - \bar{Y}_{i.k} - \bar{Y}_{.jk} + \bar{Y}_{i..} + \bar{Y}_{.j.} + \bar{Y}_{..k} - \bar{Y}_{...} \quad (3.9)$$

em que

Y_{ijk} corresponde a resposta média das b repetições do genótipo i no ambiente j e atributo k ;

$\bar{Y}_{ij.}$ é a média das respostas do i -ésimo genótipo com o j -ésimo ambiente do arranjo de interação;

$\bar{Y}_{i.k}$ é a média das respostas do i -ésimo genótipo com o k -ésimo atributo do arranjo de interação;

$\bar{Y}_{.jk}$ é a média das respostas do j -ésimo ambiente com o k -ésimo atributo do arranjo de interação;

$\bar{Y}_{i..}$ é a média das respostas do i -ésimo genótipo do arranjo de interação;

$\bar{Y}_{.j}$ é a média das respostas do j-ésimo ambiente do arranjo de interação;
 $\bar{Y}_{..k}$ é a média das respostas do k-ésimo atributo do arranjo de interação;
 $\bar{Y}_{...}$ é a média geral do experimento.

O arranjo $\underline{Z}$ de efeitos é decomposto pelo método de Tucker3 e a dimensionalidade do modelo Tucker3 foi determinada pelo método de Timmerman e Kiers (2000) apresentado na seção (2.11.2).

3.2.6 Análise de Procrustes

A estatística de Procrustes foi calculada separadamente para os genótipos e ambientes, a partir das coordenadas dos *bipLOTS*, ou seja, pelas componentes principais geradas pela análise AMMI.

O algoritmo a seguir fornece as etapas utilizadas na análise Procrustes para que a configuração $\mathbf{Y}$ corresponda à configuração $\mathbf{X}$:

1. Subtraia os vetores médios para as configurações de cada um dos respectivos pontos para ter os centróides na origem;
2. Aplique a decomposição por valores singulares na matriz $\mathbf{C} = \mathbf{Y}^T \mathbf{X}$
3. Obtenha a matriz de rotação $\mathbf{A} = \mathbf{V}\mathbf{U}^T$, em que $\mathbf{V}$ e $\mathbf{U}$ são as matrizes ortonormais geradas na decomposição por valores singulares e rotacione a configuração $\mathbf{X}$ para $\mathbf{X}\mathbf{A}$;
4. Mude a escala da configuração $\mathbf{X}$ multiplicando cada coordenada por $\hat{\rho}$, em que

$$\hat{\rho} = \text{tr}(\mathbf{X}^T \mathbf{Y} \mathbf{Y}^T \mathbf{X})^{1/2} / \text{tr}(\mathbf{X}^T \mathbf{X})$$

5. Calcular o valor da estatística M^2 de Procrustes:

$$M^2 = 1 - \frac{\{\text{tr}(\mathbf{X}^T \mathbf{Y} \mathbf{Y}^T \mathbf{X})^{1/2}\}^2}{\text{tr}(\mathbf{X}^T \mathbf{X}) \text{tr}(\mathbf{Y}^T \mathbf{Y})}$$

3.2.7 Software

Todas as análises foram executadas no *software* R Core Team (2019) (v. 3.6.0). Foram utilizados os pacotes *agricolae* para o ajuste do modelo AMMI; *rrcov3way* para o ajuste do modelo de Tucker3; *ggplot2* para a construção de todas as figuras. A dimensionalidade do modelo de Tucker3 pelo método de Timmerman e Kiers (2000) e a estatística M^2 de Procrustes foram calculadas diretamente no *software* R, sem utilizar uma função pré-definida.

4 RESULTADOS E DISCUSSÕES

4.1 Análise de dados de dupla entrada

A soma de quadrados da interação $G \times E$ e o *valor-p* do seu respectivo teste de hipótese na ANOVA univariada para cada atributo nos 6 ambientes e 19 genótipos em 3 blocos foi dada na Tabela 4.1. As pressuposições básicas da ANOVA foram satisfeitas e, portanto, conclui-se que, em todos os atributos, os efeitos da interação foram significativos, concordando com Ribeiro et al. (2004) e Spitti et al. (2019), que avaliaram a interação $G \times E$ no valor de luminosidade do feijão carioca.

Tabela 4.1. Soma de quadrados da interação genótipos $\times$ ambientes e *valor-p* do seu respectivo teste de hipótese na análise de variância para os atributos de escurecimento em grãos de feijoeiro.

Interação $G \times E$	AT1	AT2	AT3	AT4	AT5	AT6	AT7	AT8
Soma de quadrados	283,02	327,15	287,56	302,06	155,98	111,69	97,94	80,67
<i>valor-p</i>	< 0,001	< 0,001	< 0,001	< 0,001	< 0,001	< 0,001	< 0,001	0,002

Como as interações genótipos $\times$ ambientes foram significativas, prosseguiu-se com a decomposição das suas somas de quadrados, gerando matrizes $\mathbf{Y}$ de dimensão 19×6 com as médias dos $b = 3$ blocos para cada um dos oito atributos. Os principais resultados para os atributos foram apresentados a seguir.

4.1.1 Modelo AMMI para o AT1

A matriz de efeitos $\mathbf{GE}$, cujos elementos foram calculados usando (2.11), tem posto $p = \min(g - 1, e - 1) = \min(18, 5) = 5$, portanto a $SQ_{G \times E}$ poderá ser decomposta em até 5 componentes. A $SQ_{G \times E[CP_k]}$ de cada um dos 5 componentes principais no delineamento em blocos corresponde ao quadrado dos valores singulares, contidos na diagonal da matriz $\mathbf{\Lambda}$ multiplicado pelo número de bloco. Dividindo cada um desses valores pela soma de quadrados da interação, obtêm-se as proporções explicadas por cada componente principal. Tem-se, pela Tabela 4.2, que o primeiro eixo singular captura 42,88% da $SQ_{G \times E}$, o segundo 27,28%, o terceiro 14,62%, o quarto 9,65% e último 5,57%.

Atribuindo os números de graus de liberdade pelo método de Gollob, obtêm-se os quadrados médios de cada componente principal. Assim, é possível testar a significância das componentes principais, utilizando o quadrado médio do resíduo da ANOVA univariada ($QM_{Res} = 1,08$). Conclui-se a partir da Tabela 4.2 que as três primeiras componentes principais foram significativas ao nível de 5%.

Tabela 4.2. Decomposição das somas de quadrados da interação para o AT1 do escurecimento em grãos de feijoeiro.

CP	Perc. (%)	% acum.	GL	SQ	QM	F	<i>valor-p</i>
CP_1	42,88	42,88	22	121,36	5,52	5,11	< 0,001
CP_2	27,28	70,16	20	77,20	3,86	3,57	< 0,001
CP_3	14,62	84,78	18	41,38	2,30	2,13	0,0062
CP_4	9,65	94,43	16	27,32	1,71	1,58	0,0759
CP_5	5,57	100,00	14	15,76	1,13	1,05	0,4148

A Figura 4.1 apresenta o *biplot* para o modelo $AMMI_1$, com o valor médio no eixo das abscissas e CP_1 para genótipos e ambientes na ordenada. Os genótipos com maiores valores médios foram G8, G15, G5 e G12 que apresentaram valor médio de luminosidade de 54,82, 54,49, 54,31 e 54,18, respectivamente. Os genótipos G1, G2 e G4 apresentaram os menores valores, sendo iguais a 51,43, 51,90 e 51,98, respectivamente.

Figura 4.1. *Biplot* AMMI (CP_1 vs. Valores médios) para o AT1.

Considerando a proporção dos dois eixos de variabilidade explicada na Figura 4.2, que apresenta o *biplot* para as duas primeiras componentes principais que, juntas, representam 70,16% da $SQ_{G \times E}$, tem-se que o G3 é o genótipo mais estável, que apresentou um valor médio de luminosidade pouco abaixo da média. Destaca-se que os indivíduos mais estáveis são aqueles que se encontram mais próximos da origem dos eixos no gráfico do $AMMI_2$ ($CP_1 \times CP_2$). Os genótipos menos estáveis foram o genótipo G15, que possui alto valor médio de luminosidade, e o genótipo G19, com valor médio de luminosidade pouco acima da média.

Figura 4.2. *Biplot* AMMI (CP_1 vs. CP_2) para o AT1.

Considerando o fator ambiente, foi observado que os ambientes A e F (Campinas/seca e Campinas/águas) apresentaram maiores valores médios de luminosidade (55,72 e 55,39, respectivamente). Os ambientes C (Votuporanga/inverno) e E (Mococa/águas) apresentaram valores médios de luminosidade para o AT1 (53,49 e 52,83, respectivamente). E os ambientes B e D (Tatuí/seca e Ribeirão Preto/inverno) apresentaram os menores valores médios de luminosidade (51,28 e 50,79) (Figura 4.1). Pela Figura 4.2

tem-se que o ambiente B (Tatuí/seca) é o mais estável (mais próximo de zero em relação às duas primeiras componentes principais) e o ambiente F (Campinas/águas) é o menos estável.

Em relação às adaptações específicas, tem-se pela Figura 4.2, que o genótipo G15 tem uma interação positiva com o ambiente F e negativa como ambiente D, o genótipo G19 tem uma interação positiva com o ambiente A e negativa como ambiente F e o genótipo G9 tem uma interação positiva com o ambiente E e negativa como ambiente C.

Pelo teste F (Tabela 4.3), as três primeiras componentes principais são significativas e devem ser analisadas conjuntamente para concluir sobre a estabilidade dos genótipos e ambientes, entretanto as conclusões considerando somente CP_1 e CP_2 explicam a maior parte da $SQ_{G \times E}$ (70,16%).

4.1.2 Modelo AMMI para o AT2

Para o AT2 (Tabela 4.3), as três primeiras componente principais foram significativas ao nível de 5%, considerando um $QM_{Res} = 1,24$. A primeira componente explicou 47,00% da $SQ_{G \times E}$, a CP_2 explicou 23,80 e a CP_3 , 16,80%, explicando juntas 87,60% da soma de quadrados da interação.

Tabela 4.3. Decomposição das somas de quadrados da interação para o AT2 do escurecimento em grãos de feijoeiro.

CP	Perc. (%)	% acum.	GL	SQ	QM	F	valor-p
CP_1	47,00	47,00	22	153,63	6,98	5,63	< 0,001
CP_2	23,80	70,80	20	77,91	3,90	3,14	< 0,001
CP_3	16,80	87,60	18	55,02	3,06	2,46	0,0012
CP_4	7,70	95,30	16	25,28	1,58	1,27	0,2184
CP_5	4,70	100,00	14	15,31	1,09	0,88	0,5814

Figura 4.3. Biplot AMMI (CP_1 vs. Valores médios) para o AT2.

Pelas Figuras 4.3 e 4.4, tem-se que os genótipos G8, G5, G15 e G12 apresentam maiores valores de luminosidade, sendo iguais a 53,04, 52,66, 52,61 e 52,41 na variável AT2, respectivamente, sendo o G12 um genótipo com alta estabilidade. Os genótipos G1, G2 e G4 apresentaram os menores valores de luminosidade no AT2 (49,57, 50,01 e 50,03, respectivamente). Os genótipos mais estáveis foram G12 e G11 e o menos estável foi o genótipo G15.

Os ambientes que apresentaram maiores valores médios de luminosidade foram os ambientes F (Campinas/águas) e A (Campinas/seca), sendo que o ambiente F é ambiente menos estável. Os ambientes B (Tatuí/seca) e D (Ribeirão Preto/inverno) apresentaram menores valores médios de luminosidade para o AT2, sendo que o ambiente B apresentou maior estabilidade.

Em relação às adaptações específicas, o genótipo G15 exibiu uma interação positiva com o ambiente F enquanto que o genótipo G19 exibiu uma interação positiva com o ambiente C e negativa como ambiente E. O genótipo G4 teve uma interação positiva com o ambiente D (Figura 4.4).

Figura 4.4. Biplot AMMI (CP_1 vs. CP_2) para o AT2.

4.1.3 Modelo AMMI para o AT3

A Tabela 4.4 apresenta a decomposição das somas de quadrados da interação para o AT3 considerando um $QM_{Res} = 1,32$. A primeira componente principal explicou 47,50% da $SQ_{G \times E}$, a segunda componente explicou 25,60% e a terceira, 14,30%. As três primeiras componentes foram significativas ao nível de 5% e explicaram juntas 87,40% da soma de quadrados da interação.

Tabela 4.4. Decomposição das somas de quadrados da interação para o AT3 do escurecimento em grãos de feijoeiro.

CP	Perc. (%)	% acum.	GL	SQ	QM	F	valor-p
CP_1	47,50	47,50	22	136,70	6,21	4,70	< 0,001
CP_2	25,60	73,20	20	73,65	3,68	2,78	< 0,001
CP_3	14,30	87,40	18	41,12	2,28	1,73	0,0360
CP_4	7,80	95,20	16	22,38	1,40	1,06	0,3954
CP_5	4,80	100,00	14	13,71	0,98	0,74	0,7326

Para o atributo AT3 (Figuras 4.5 e 4.6), os genótipos G8, G15, G5 e G12, foram os que apresentam maiores valores de luminosidade (51,43, 51,07, 50,92 e 50,84, respectivamente), sendo entre eles o G12 o genótipo com maior estabilidade. Os genótipos G1, G2 e G4 (48,50, 48,52 e 48,54, respectivamente) apresentaram os menores valores de luminosidade neste atributo. O genótipo mais estável foi o G3.

Observa-se ainda que os ambientes que apresentaram maiores valores médios de luminosidade foram os ambientes F (Campinas/águas) e A (Campinas/seca), sendo o ambiente F o ambiente que

exibiu menor estabilidade. O ambiente Tatuí/seca (B) foi o mais estável, considerando o gráfico *biplot* da primeira e segunda componentes principais, explicando 73,20% da $SQ_{G \times E}$.

Em relação às adaptações específicas, o genótipo G15 exibiu uma interação positiva com o ambiente F. Os genótipos G2 e G19 exibiram uma interação positiva com o ambiente C e negativa como ambiente E.

Figura 4.5. *Biplot* AMMI (CP_1 vs. Valores médios) para o AT3.

Figura 4.6. *Biplot* AMMI (CP_1 vs. CP_2) para o AT3.

4.1.4 Modelo AMMI para o AT4

A Tabela 4.5 indica que as duas primeiras componentes principais foram significativas ao nível de 5% para o AT4, considerando $QM_{Res} = 1,27$. Juntas elas explicaram um total de 73,30% da $SQ_{G \times E}$.

Tabela 4.5. Decomposição das somas de quadrados da interação para o AT4 do escurecimento em grãos de feijoeiro.

CP	Perc. (%)	% acum.	GL	SQ	QM	F	valor-p
CP_1	44,00	44,00	22	133,01	6,05	4,76	< 0,001
CP_2	29,30	73,30	20	88,46	4,42	3,48	< 0,001
CP_3	12,30	85,60	18	37,08	2,06	1,62	0,0570
CP_4	7,90	93,50	16	23,79	1,49	1,17	0,2940
CP_5	6,50	100,00	14	19,73	1,41	1,11	0,3502

Para o atributo AT4 (Figura 4.7 e Figura 4.8), os genótipos G8, G5, G15 e G12 apresentam maiores valores de luminosidade. Os genótipos G1, G2, G4, G18 e G10 apresentaram os menores valores de luminosidade. O genótipo mais estável foi o G3 e o genótipo menos estável foi o G15.

Figura 4.7. Biplot AMMI (CP_1 vs. Valores médios) para o AT4.

Figura 4.8. Biplot AMMI (CP_1 vs. CP_2) para o AT4.

Os ambientes que apresentaram maiores valores médios de luminosidade foram os ambientes A (Campinas/seca) e F (Campinas/águas). Os ambientes D (Ribeirão Preto/inverno) e B (Tatuí/seca) apresentaram os menores valores médios de luminosidade. O ambiente Campinas/seca foi o que apresentou maior estabilidade e o ambiente D (Ribeirão Preto/inverno) exibiu menor estabilidade no AT4.

Em relação às adaptações específicas, os genótipos G15, G7, G10 e G13 exibiram uma interação positiva com o ambiente F e negativa com o ambiente D. Os genótipos G2 e G19 exibiram uma interação positiva com o ambiente C e negativa com o ambiente E. Os genótipos G2 e G19 exibiram uma interação positiva com o ambiente C e negativa com o ambiente E (Figura 4.8).

4.1.5 Modelo AMMI para o AT5

Para o atributo AT5 do estudo (Tabela 4.6), as três primeiras componente principais foram significativas ao nível de 5%, considerando $QM_{Res} = 0,94$. A primeira componente explicou 43,60% da $SQ_{G \times E}$, a CP_2 explicou 30,30% e a CP_3 18,30%, explicando juntas 92,20%.

Tabela 4.6. Decomposição das somas de quadrados da interação para o AT5 do escurecimento em grãos de feijoeiro.

CP	Perc. (%)	% acum.	GL	SQ	QM	F	valor-p
CP_1	43,60	43,60	22	67,97	3,09	3,29	< 0,001
CP_2	30,30	73,90	20	47,25	2,36	2,51	< 0,001
CP_3	18,30	92,20	18	28,51	1,58	1,68	0,0427
CP_4	4,30	96,50	16	6,77	0,42	0,45	0,9668
CP_5	3,50	100,00	14	5,47	0,39	0,41	0,9675

Os genótipos G15, G5, G6 e G8 apresentaram os maiores valores médios de luminosidade, sendo iguais a 49,34, 49,25, 49,21 e 49,20, respectivamente. Os genótipos G15, G12 e G11 apresentaram alta estabilidade. Os genótipos G2, G18 e G1 apresentaram os menores valores médios de luminosidade no AT5 (47,33, 47,40 e 47,58, respectivamente), sendo o G2 o genótipo menos estável (Figuras 4.9 e 4.10).

Figura 4.9. Biplot AMMI (CP_1 vs. Valores médios) para o AT5.

Pode-se observar ainda pelas Figuras 4.9 e 4.10 que os ambientes que apresentaram maiores valores médios de luminosidade foram os ambientes A (Campinas/seca) e F (Campinas/águas). Os ambientes D (Ribeirão Preto/inverno) e B (Tatuí/seca) apresentaram menores valores médios de luminosidade

para o AT5. O ambiente mais estável foi o ambiente A (Campinas/seca) e o menos estável foi o ambiente C (Votuporanga/inverno).

Em relação às adaptações específicas, tem-se pela Figura 4.10, que o genótipo G14 exibiu uma interação positiva com o ambiente D e negativa com o ambiente B. O genótipo G10 exibiu uma interação positiva com o ambiente B. E os genótipos G2 e G19 exibiram uma interação negativa como ambiente E.

Figura 4.10. Biplot AMMI (CP_1 vs. CP_2) para o AT5.

4.1.6 Modelo AMMI para o AT6

A Tabela 4.7 apresenta a decomposição das somas de quadrados da interação para o AT6 e mostra que a primeira componente principal explicou 39,70% da soma de quadrados da interação e a segunda componente explicou 31,90%. As duas primeiras componentes foram significativas ao nível de 5%, considerando $QM_{Res} = 0,73$ e explicaram juntas 71,60% da $SQ_{G \times E}$.

Tabela 4.7. Decomposição das somas de quadrados da interação para o AT6 do escurecimento em grãos de feijoeiro.

CP	Perc. (%)	% acum.	GL	SQ	QM	F	valor-p
CP_1	39,70	39,70	22	44,38	2,02	2,77	< 0,001
CP_2	31,90	71,60	20	35,68	1,78	2,44	< 0,001
CP_3	13,20	84,80	18	14,77	0,82	1,12	0,3246
CP_4	9,80	94,60	16	10,89	0,68	0,93	0,5243
CP_5	5,40	100,00	14	5,96	0,43	0,59	0,8715

Para o atributo AT6, o G8 foi o que apresentou maior valor médio de luminosidade (47,31). Os genótipos G1, G2 e G18 (45,17, 45,38 e 45,41, respectivamente) apresentaram os menores valores de luminosidade. O genótipo menos estável foi o genótipo G2 e nenhum genótipo foi considerado estável nesse atributo (Figuras 4.11 e 4.12).

Os ambientes que apresentaram maiores valores médios de luminosidade foram os ambientes A (Campinas/seca) e F (Campinas/águas). Os ambientes D (Ribeirão Preto/inverno) e B (Tatuí/seca) foram os que tiveram menores valores médios de luminosidade. Nesse atributo nenhum ambiente apresentou grande estabilidade. O ambiente Mococa/águas (E) exibiu a menor estabilidade.

Em relação às adaptações específicas, o genótipo G9 exibiu uma interação positiva com o ambiente E, enquanto que o genótipo G13 exibiu uma interação negativa neste mesmo ambiente.

Figura 4.11. Biplot AMMI (CP_1 vs. Valores médios) para o AT6.

Figura 4.12. Biplot AMMI (CP_1 vs. CP_2) para o AT6.

4.1.7 Modelo AMMI para o AT7

Para o AT7, apenas as duas primeiras componentes principais foram significativas ao nível de 5%, considerando $QM_{Res} = 0,61$. Juntas elas explicaram um total de 75,20% da $SQ_{G \times E}$ (Tabela 4.8).

Tabela 4.8. Decomposição das somas de quadrados da interação para o AT7 do escurecimento em grãos de feijoeiro.

CP	Perc. (%)	% acum.	GL	SQ	QM	F	valor-p
CP_1	41,30	41,30	22	40,41	1,84	3,02	< 0,001
CP_2	33,90	75,20	20	33,20	1,66	2,72	< 0,001
CP_3	13,60	88,80	18	13,29	0,74	1,21	0,2467
CP_4	6,60	95,40	16	6,46	0,40	0,66	0,8218
CP_5	4,60	100,00	14	4,58	0,33	0,54	0,9074

Figura 4.13. Biplot AMMI (CP_1 vs. Valores médios) para o AT7.

Figura 4.14. Biplot AMMI (CP_1 vs. CP_2) para o AT7.

Para o atributo AT7 (Figura 4.13 e Figura 4.14), o genótipo G8 apresentou maior valor médio de luminosidade (45,69). Os genótipos G1, G2, e G18 (43,85, 43,94 e 43,94) apresentaram os menores valores de luminosidade. O genótipo mais estável foi o G11 e os genótipos menos estáveis foram o G2 e G19.

Os ambientes que apresentaram maiores valores médios de luminosidade foram os ambientes A e F. Os ambientes B (Tatuí/seca) e D (Ribeirão Preto/inverno) apresentaram os menores valores médios de luminosidade. Nenhum ambiente foi considerado estável nesse atributo. O ambiente C (Votuporanga/inverno) foi o menos estável do AT7. Em relação às adaptações específicas, o genótipo G19 apresentou uma interação positiva com o ambiente C (Figura 4.14).

4.1.8 Modelo AMMI para o AT8

Para o atributo AT8 (Tabela 4.9), as duas primeiras componente principais foram significativas ao nível de 5%, considerando $QM_{Res} = 0,55$. A primeira componente explicou 47,80% da $SQ_{G \times E}$ e a CP_2 explicou 28,10%, assim, as duas primeiras componentes principais explicam juntas 75,90%.

Tabela 4.9. Decomposição das somas de quadrados da interação para o AT8 do escurecimento em grãos de feijoeiro.

CP	Perc. (%)	% acum.	GL	SQ	QM	F	valor-p
CP_1	47,80	47,80	22	38,57	1,75	3,18	< 0,001
CP_2	28,10	75,90	20	22,64	1,13	2,05	0,0059
CP_3	11,20	87,10	18	9,04	0,50	0,91	0,5550
CP_4	7,40	94,50	16	6,00	0,38	0,69	0,8025
CP_5	5,50	100,00	14	4,42	0,32	0,58	0,8792

Figura 4.15. Biplot AMMI (CP_1 vs. Valores médios) para o AT8.

O genótipo G8 apresentou o maior valor médio de luminosidade, sendo igual a 44,57. Os genótipos G1, G18 e G2 apresentaram os menores valores de luminosidade no AT8 (42,86, 42,87 e 43,06, respectivamente). O genótipo mais estável foi G17 e o menos estável foi o G2 (Figuras 4.15 e 4.16).

O ambiente A (Campinas/seca) foi o que apresentou o maior valor médio de luminosidade e o ambiente B (Tatuí/seca) apresentou o menor valor médio de luminosidade. Nenhum ambiente foi considerado estável no AT8. O ambiente E (Mococa/águas) foi o que apresentou a menor estabilidade. Em relação às adaptações específicas, o genótipo G7 exibiu uma interação positiva com o ambiente F e negativa com o ambiente D.

Figura 4.16. *Biplot* AMMI (CP_1 vs. CP_2) para o AT8.

4.1.9 Conclusões gerais sobre os resultados das análises AMMI individuais

Conclui-se que o genótipo G8 apresentou o maior valor médio de luminosidade em todos os atributos analisados, com exceção ao atributo AT5, no qual os genótipos que apresentaram os maiores valores médios de luminosidade foram o G15, G5 e G6, seguido pelo G8. Nos atributos AT1, AT2, AT3 e AT4, os quatro genótipos com maiores valores médios de luminosidade foram G8, G15, G5 e G12.

Os genótipos G1 e G2 apareceram entre aqueles com menores valores médios de luminosidade em todos os oito atributos analisados. O G4 apareceu entre aqueles com menor valor médio de luminosidade nos quatro primeiros atributos e nos atributos AT5, AT6, AT7 e AT8, o genótipo G18 foi o que apareceu entre os menores valores médios de luminosidade.

Em todos os atributos analisados, os ambientes A (Campinas/seca) e F (Campinas/águas) foram os que apresentaram os maiores valores de luminosidade. Os ambientes B (Tatuí/seca) e D (Ribeirão Preto/inverno), foram aqueles que apresentaram os menores valores médios de luminosidade. Conclui-se que os efeitos genéticos contribuíram mais para a interação $G \times E$ do que os efeitos ambientais, concordando com os resultados obtidos por Spitti et al. (2019), que avaliaram os resultados do mesmo experimento.

Em relação a estabilidade, o G3 foi considerado entre os mais estáveis em quatro dos oito atributos e (AT1, AT2, AT3 e AT4) e o genótipo G11 foi considerado entre os mais estáveis em três dos oito atributos (AT2, AT5 e AT7). Nos atributos AT2 e AT5, o G12 foi considerado estável e no atributo AT8, o genótipo G17 foi considerado estável. No AT5, o genótipo G15 foi considerado o mais estável, porém este genótipo foi considerado menos estável no AT1, AT2, AT3 e AT4. Nos atributos AT5, AT6, AT7 e AT8 o genótipo G2 foi o menos estável e nos AT1 e AT7, o genótipo G19 apareceu entre os menos estáveis.

Os ambientes mais estáveis foram: o ambiente B (Tatuí/seca) no AT1, AT2 e AT3 e o ambiente o A (Campinas/seca) nos atributos AT4 e AT5. No AT6, AT7 e AT8. Nos demais atributos, nenhum ambiente foi considerado estável. Já os ambientes que apresentaram menor estabilidade foram: o ambiente F (Campinas/águas) no AT1, AT2 e AT3, o ambiente D (Ribeirão Preto/inverno) no AT4, o ambiente C (Votuporanga/inverno) nos atributos AT5 e AT7 e o ambiente E (Mococa/águas) nos atributos AT6 e AT8.

Assim, conclui-se que o ambiente F (Campinas/águas) apresentou alto valor médio de lumino-

sidade e baixa estabilidade e o ambiente B (Tatuí/seca) apresentou em grande parte dos atributos baixo valor médio de luminosidade e alta estabilidade. O ambiente A (Campinas/seca) possui as características mais desejadas (alto valor médio de luminosidade e alta estabilidade) apenas nos atributos AT6 e AT7.

O genótipo G15 apresentou alto valor médio de luminosidade e baixa estabilidade e o genótipo G8 apresentou alto valor médio de luminosidade, sendo considerado estável em grande parte dos atributos. Ribeiro et al. (2014) indicam que os genótipos com maiores valores no mercado apresentam valor de luminosidade superior a 55, pois o consumidor associa o escurecimento dos grãos com lento cozimento, porém nenhum genótipo do experimento apresentou valor médio de luminosidade superior ao desejado.

Entretanto, Spitti et al. (2019) sugeriram que se utilize como critério o valor de luminosidade igual a 53 no momento da colheita, uma vez que, abaixo desse valor o mercado refuta o produto, no entanto há maior tolerância a doenças e melhor qualidade do caldo (caldo grosso). Acima desse valor, o produto apresenta maior aceitação no mercado, porém é mais suscetível a doenças do solo e apresenta uma qualidade mais baixa de caldo (caldo claro). Assim, considerando o AT1, os genótipos G8 e G15 apresentam conformidade com o valor desejado de luminosidade.

4.2 Análise de dados de tripla entrada

A análise de variância para três fatores, considerando os 8 atributos em 19 genótipos e 6 ambientes em um delineamento experimental aleatorizado em 3 blocos sobre o escurecimento em grãos de feijoeiro indicou que a interação tripla não foi significativa ($\text{valor-}p = 0,9959$), de forma que não é necessária a decomposição da interação tripla. O pressuposto de homogeneidade de variância não foi atendido, portanto foi realizada a transformação de dados sugerida pelo método de Box-Cox, na qual se utilizou $\lambda = 4$.

Uma interação tripla não significativa implica que as conclusões feitas para o modelo $G \times E$ são suficientes e o modelo Tucker3 apresentará resultados semelhantes aos obtidos para a análise de dados de dupla entrada. Porém, o desdobramento da interação tripla foi executado de forma a confirmar as conclusões dos resultados por atributos, para isso gerou-se o arranjo cúbico $\underline{Z}$ de dimensão $(19 \times 6 \times 8)$ com o efeito das interações triplas, obtido como em (3.9). A decomposição do arranjo $\underline{Z}$ foi feita por meio do modelo Tucker3.

4.2.1 Ajuste do modelo Tucker3

Para selecionar o melhor modelo Tucker3, utilizou-se o método de Timmerman e Kiers (2000). Todos os modelos satisfizeram as condições de $P \leq QR = 48$, $Q \leq PR = 152$, $R \leq PQ = 114$. Considerando os modelos com o mesmo número total de componentes S , selecionou-se aquele com a maior proporção da soma de quadrados ajustada (SQ_S).

Para comparar modelos com número total de componentes diferente, calculou-se a quantidade $\text{dif}_S = SQ_S - SQ_{S-1}$. Para a solução trivial $P = 1$, $Q = 1$ e $R = 1$, Schepers et al. (2008) sugerem que se utilize $\text{dif}_S = 0$. O ponto de corte para os dif_S foi igual a 3,80. Os modelos com dif_S maiores que 3,80 foram apresentados na Tabela 4.10.

Tabela 4.10. Resultados do procedimento de Timmerman-Kiers para selecionar o modelo de Tucker3 para o experimento sobre o escurecimento em grãos de feijoeiro.

P	Q	R	S	SQ_S	dif_S	b_S
2	2	1	5	50,15	12,47	2,10
3	3	1	7	58,03	5,93	1,29
4	3	2	9	66,18	4,58	0,98

Assim, todo valor de $diff_S$ menor que 3,80 foi desconsiderado no cálculo do b_S e a solução escolhida foi o modelo Tucker3(2,2,1), que explica 50,15% da soma de quadrados da interação tripla entre genótipos $\times$ ambientes $\times$ atributos. Sendo este percentual de variabilidade explicada menor do que o obtido por Varela et al. (2006) e Araújo et al. (2010), que utilizaram, respectivamente, quatro e três atributos. Porém, maior do que o obtido por Araújo et al. (2011), que utilizou 10 atributos.

As matrizes $\mathbf{A}$, $\mathbf{B}$ e $\mathbf{C}$ e o arranjo núcleo $\mathbf{G}$ foram apresentadas na Tabela 4.11. As duas primeiras componentes, p_1 e p_2 , da matriz $\mathbf{A}$ (referente a genótipos) explicam 37,57% e 12,58%, respectivamente. Estes valores foram obtidos por meio da equação (3.4).

As duas componentes q_1 e q_2 da matriz $\mathbf{B}$ (que se refere aos ambientes) também explicam 37,57% e 12,58%, respectivamente e a matriz $\mathbf{C}$, que é formada apenas por uma componente, r_1 , explica 50,15% da $SQ_{G \times E \times A}$. Estes valores foram obtidos por meio das equações (3.5) e (3.6).

O arranjo núcleo $\mathbf{G}(2, 2, 1)$ mostra a relação entre as componentes e indica que a relação mais importante é entre as primeiras componentes de cada fator, $g_{111} = -943561,4661$. Por meio da equação (3.7), este valor indica que a combinação da primeira componente dos genótipos com a primeira componente dos ambientes e com a única componente dos atributos explicam juntas 37,58% da $SQ_{G \times E \times A}$. A segunda relação mais importante é entre as segundas componentes dos genótipos e dos ambientes com a única componente dos atributos, $g_{221} = 45954,3465$, que explica 12,58% da soma de quadrados da interação tripla entre genótipos $\times$ ambientes $\times$ atributos. As outras duas relações g_{121} e g_{211} apresentam valores menores e são as relações menos importantes apresentadas em $\mathbf{G}$.

Tabela 4.11. Escores das componentes principais para um modelo de Tucker3(2,2,1) para o arranjo da interação tripla entre genótipos $\times$ ambientes $\times$ atributos.

Genótipos (<i>A</i>)			Ambientes (<i>B</i>)			Atributos (<i>C</i>)	
	<i>p</i> ₁	<i>p</i> ₂		<i>q</i> ₁	<i>q</i> ₂		<i>r</i> ₁
G1	-0,2192	-0,4598	A	-0,2322	0,2832	AT1	-0,4548
G2	-0,2086	-0,2629	B	-0,0876	-0,0652	AT2	-0,4586
G3	-0,0319	-0,0908	C	-0,1440	0,7046	AT3	-0,2844
G4	-0,2156	0,2146	D	-0,4669	-0,6049	AT4	-0,1534
G5	-0,1790	0,4000	E	0,1001	-0,1958	AT5	0,2208
G6	-0,1372	-0,0548	F	0,8305	-0,1220	AT6	0,3353
G7	0,2306	-0,0173				AT7	0,3725
G8	0,2379	0,1541				AT8	0,4225
G9	0,0302	0,2904					
G10	0,2514	-0,0067					
G11	0,0382	-0,0218					
G12	0,0493	-0,1202					
G13	0,2363	-0,0997					
G14	0,0050	0,1578					
G15	0,5916	-0,2275					
G16	-0,0775	-0,0390					
G17	-0,3755	0,1072					
G18	0,0514	0,4099					
G19	-0,2771	-0,3335					

Arranjo Núcleo (<i>G</i>)			
<i>r</i> ₁			
	<i>q</i> ₁	<i>q</i> ₂	
<i>p</i> ₁	-943561,4661	5,0436	
<i>p</i> ₂	-23,8132	545954,3465	

Os resultados também mostraram que a matriz $\mathbf{C}$ foi caracterizada por um contraste entre AT1, AT2, AT3 e AT4 (apresentando escores negativos) e AT5, AT6, AT7 e AT8 (escores positivos). Convém notar que o contraste desta componente é dado entre os atributos avaliados pelo método de escurecimento natural e pelo método de escurecimento acelerado.

O *joint plot* da Figura 4.17 corresponde ao *biplot* da matriz $\mathbf{\Delta}_1 = \mathbf{A}\mathbf{G}_1\mathbf{B}^T$, de forma que $\mathbf{G}_1$ é a fatia frontal de $\mathbf{G}$ do modelo Tucker3(2,2,1). Este *joint plot* é projetado dentro da componente

r_1 que explica 50,15% da $SQ_{G \times E \times A}$ e apresenta a interação entre genótipos e ambientes nos atributos avaliados pelo método de escurecimento natural e pelo método de escurecimento acelerado. Desta forma, em relação aos atributos avaliados pelo método de escurecimento natural do grão (AT1, AT2, AT3 e AT4), observa-se que:

- o genótipo G17 teve uma interação negativa com o ambiente F (Campinas/águas) e positiva com o ambiente D (Ribeirão Preto/inverno);
- o genótipo G15 teve uma interação negativa com o ambiente D e positiva com o ambiente F;
- o genótipo G18 teve uma interação negativa com o ambiente C (Votuporanga/inverno) e uma interação positiva com o ambiente E (Mococa/águas);
- o genótipo G8 teve uma interação negativa com o ambiente A (Campinas/seca) e positiva com o ambiente E;
- os genótipos G1, G2 e G19 tiveram uma interação negativa com o ambiente E e positiva com o ambiente A e C;
- o genótipo G4 apresentou uma interação positiva com o ambiente D.

Figura 4.17. Joint plot projetado dentro da primeira componente do terceiro modo para o modelo Tucker3(2,2,1).

Para os atributos avaliados pelo método de escurecimento acelerado do grão (AT5, AT6, AT7 e AT8), as interpretações são opostas as observadas anteriormente, isto é:

- o genótipo G17 teve uma interação positiva com o ambiente F (Campinas/águas) e negativa com o ambiente D (Ribeirão Preto/inverno);
- o genótipo G15 teve uma interação positiva com o ambiente D e negativa com o ambiente F;
- o genótipo G18 teve uma interação positiva com o ambiente C (Votuporanga/inverno) e uma interação negativa com o ambiente E (Mococa/águas);
- o genótipo G8 teve uma interação positiva com o ambiente A (Campinas/seca) e negativa com o ambiente E;

- os genótipos G1, G2 e G19 tiveram uma interação positiva com o ambiente E e negativa com o ambiente A e C;
- o genótipo G4 apresentou uma interação negativa com o ambiente D.

Observa-se ainda que os genótipos G11, G12 e G3, localizados na região central do *biplot* tiveram desempenho médio em todos os atributos avaliados e que os ambientes C, D e F (setas longas em relação ao centro do *biplot*) representaram mais a variabilidade da interação genótipos $\times$ ambientes associada ao componente do terceiro modo (atributos) do que os ambientes A, B e E (setas curtas).

Os resultados mostraram que os genótipos G3, G11 e G12, que foram os genótipos mais estáveis em grande parte dos atributos avaliados individualmente, também são tidos como estáveis na Figura 4.17. O genótipo G17 apresentou um comportamento diferente do observado para o AMMI individual AT8, uma vez que foi considerado estável neste atributo. O genótipo G15 foi considerado um genótipo estável no AT5 e instável em todos os atributos avaliados pelo método de escurecimento natural do grão (AT1, AT2, AT3 e AT4), sendo um genótipo de destaque na Figura 4.17 pela sua distância em relação ao centro do *biplot*. O genótipo G2 apresentou comportamento instável em todos os atributos avaliados pelo método de escurecimento acelerado do grão (AT5, AT6, AT7 e AT8), estando razoavelmente longe do centro no *joint plot* e o genótipo G19, que foi considerado instável em AT1 e AT6, também apresentou comportamento similar ao G2 avaliando-se a Figura 4.17.

O ambiente B (Tatuí/seca), considerado estável pelo *joint plot*, também foi estável em AT1, AT2 e AT3 nas análises AMMI feitas para interação $G \times E$. O ambiente E (Mococa/águas), também estável por Tucker3, foi um ambiente instável em AT6 e AT8. O ambiente F (Campinas/águas) foi o menos estável nos atributos AT1, AT2 e AT3 e o ambiente D (Ribeirão Preto/inverno) foi o menos estável em AT4, sendo estes dois ambientes os menos estáveis pela Figura 4.17.

Spitti et al. (2019), que analisaram os resultados do mesmo experimento por meio da análise de variância multivariada (MANOVA) e análise de contrastes entre genótipos, indicaram que o genótipo BRS Perola (G1) apresentou um contraste significativo com os genótipos G18 e G8. Avaliando-se a Figura 4.17, estes contrastes ficam evidentes, uma vez que estes genótipos estão dispostos em oposição no *joint plot*. Os autores indicaram também um contraste entre os genótipos G1 e G17, a Figura 4.17 exibiu um comportamento heterogêneo entre os genótipos, mesmo não estando em oposição.

4.2.2 Análise de Procrustes

A estatística de Procrustes foi calculada com a finalidade de comparar os *biplots* da análise AMMI individual com o *joint plot* do modelo Tucker3(2,2,1).

O *joint plot* da análise Tucker3(2,2,1) e os *biplots* para os atributos tiveram valores de M^2 similares após a transformação de procrustes para o fator genótipo. O AT2 apresentou um valor menor da estatística, isto é, ele foi o mais semelhante ao *joint plot* em relação aos genótipos. Considerando os ambientes, os *biplots* do AT1, AT2 e AT3 foram os mais similares ao *joint plot* da análise Tucker3(2,2,1), enquanto que os *biplots* dos atributos AT5, AT6 e AT8 exibiram menor similaridade ao *joint plot* (Tabela 4.12).

O valor da estatística M^2 de Procrustes é maior para os genótipos por conter maior número de observações do que os ambientes. Desta forma, essa estatística não deve ser usada para comparar os genótipos e os ambientes, não sendo correto afirmar que os genótipos apresentaram maior similaridade do que os ambientes. O que é possível afirmar é que, após as transformações de procrustes, os *biplots* do AT1, AT2 e AT3 ficaram mais similares ao *joint plot* da análise Tucker3(2,2,1) em relação aos ambientes.

Tabela 4.12. Estatística de Procrustes (M^2).

	Genótipos	Ambientes
<i>biplot</i> AT1 $\times$ <i>joint plot</i>	0,9329	0,1073
<i>biplot</i> AT2 $\times$ <i>joint plot</i>	0,8854	0,1043
<i>biplot</i> AT3 $\times$ <i>joint plot</i>	0,9243	0,1269
<i>biplot</i> AT4 $\times$ <i>joint plot</i>	0,9285	0,2392
<i>biplot</i> AT5 $\times$ <i>joint plot</i>	0,9860	0,4916
<i>biplot</i> AT6 $\times$ <i>joint plot</i>	0,9689	0,4963
<i>biplot</i> AT7 $\times$ <i>joint plot</i>	0,9827	0,3909
<i>biplot</i> AT8 $\times$ <i>joint plot</i>	0,9293	0,4788

Analisando os *biplots* desses 3 atributos e o *joint plots* de Tucker3, observa-se que em todos eles o ambiente B foi o mais estável e o ambiente F foi o menos estável, evidenciando o resultado obtido pela análise de Procrustes. Além disso, em relação às adaptações específicas, houve mais conformidades entre o *joint plot* e as análises individuais dos quatro primeiros atributos, referente ao método de escurecimento natural, do que entre a análise de Tucker3 e os atributos AT5, AT6, AT7 e AT8.

4.3 Análise de dados de tripla entrada no caso bivariado

Para comparar os dois métodos de escurecimento de grãos de feijão, um novo ajuste do modelo de Tucker3 foi feito para os últimos atributos de cada método (AT4 e AT8), que correspondem ao valor de luminosidade dos grãos avaliados aos 90 dias e armazenados em condições de prateleira com temperatura e umidade ambiente (escurecimento natural) e às 96 horas dos grãos armazenados em condições de câmara com luz ultravioleta e fluorescente (escurecimento acelerado), respectivamente.

A análise de variância para três fatores, considerando os atributos AT4 e AT8 em 19 genótipos e 6 ambientes em 3 blocos, indicou que a interação tripla foi significativa considerando $\alpha = 0,10$ (*valor-p* = 0,0934), portanto, prosseguiu-se com o ajuste do modelo de Tucker3.

4.3.1 Ajuste do modelo Tucker3

A seleção do modelo de Tucker3 foi feita pelo método de Timmerman-Kiers. Inicialmente, selecionaram-se os modelos que satisfazem as condições de $P \leq QR = 12$, $Q \leq PR = 38$, $R \leq PQ = 114$, como $P = 19 > 12$, foram utilizados apenas os modelos até $P = 12$. Considerando, agora, os modelos com o mesmo número total de componentes S , selecionou-se aquele com a maior proporção da soma de quadrados ajustada (SQ_S).

Tabela 4.13. Resultados do procedimento de Timmerman-Kiers para selecionar o modelo de Tucker3 para o experimento sobre o escurecimento em grãos de feijoeiro no caso bivariado.

P	Q	R	S	SQ_S	$diff_S$	b_S
2	2	1	5	70,12	21,63	1,69
3	3	1	7	82,91	12,79	1,16
4	4	1	9	93,90	10,99	1,80

Para comparar modelos com número total de componentes diferente, calculou-se a quantidade $diff_S = SQ_S - SQ_{S-1}$. Os modelos $diff_S$ maiores que o ponto de corte (1,40) foram apresentados na Tabela 4.13. Após o cálculo do b_S , a solução escolhida foi o modelo Tucker3(4,4,1), que explicou 93,90% da soma de quadrados da interação tripla. A porcentagem da variabilidade explicada por este modelo foi maior que no modelo selecionado por Araújo et al. (2010) em um estudo sobre a produtividade de 13 genótipos de feijão, no qual o modelo selecionado explicou 90,38%.

Tabela 4.14. Escores das componentes principais para o modelo de Tucker3(4,4,1) para as matrizes **A**, **B** e **C**.

Genótipos (<i>A</i>)					Ambientes (<i>B</i>)				
	<i>p</i> ₁	<i>p</i> ₂	<i>p</i> ₃	<i>p</i> ₄		<i>q</i> ₁	<i>q</i> ₂	<i>q</i> ₃	<i>q</i> ₄
G1	-0,0446	-0,2557	0,2637	0,1603	A	-0,0531	0,4390	0,7528	0,1099
G2	0,0477	-0,3364	0,5603	0,2833	B	0,0582	-0,2055	0,2049	-0,4731
G3	-0,0591	0,0131	-0,0449	-0,0061	C	-0,0856	0,6687	-0,5685	0,1175
G4	-0,2621	-0,0098	-0,1206	-0,0653	D	0,7440	-0,1470	-0,1875	-0,2154
G5	-0,2818	0,3155	-0,1574	-0,0651	E	-0,0056	-0,4522	-0,0212	0,7772
G6	-0,1688	-0,1793	0,2799	-0,2859	F	-0,6580	-0,3029	-0,1805	-0,3161
G7	0,2100	0,0715	-0,0928	0,3341					
G8	0,0116	0,3269	0,2912	-0,3180					
G9	-0,2325	0,2252	0,1495	-0,1376					
G10	0,2517	0,0645	-0,2961	0,3639					
G11	0,0736	-0,0006	-0,1528	0,0230					
G12	0,1201	-0,1288	-0,1740	-0,1865					
G13	0,3023	0,2374	0,1327	0,0991					
G14	-0,0659	0,0719	-0,1185	-0,3218					
G15	0,5912	-0,0151	-0,0526	-0,3328					
G16	-0,1364	-0,1604	0,0890	0,0975					
G17	-0,4055	-0,1471	-0,3207	0,2000					
G18	-0,0243	0,3967	0,0648	0,3221					
G19	0,0729	-0,4895	-0,3006	-0,1643					

Atributos (<i>C</i>)	
	<i>r</i> ₁
AT4	-0,7071
AT8	0,7071

As quatro componentes, p_1 , p_2 , p_3 e p_4 , da matriz **A** (referente a genótipos) explicam 48,49%, 21,63%, 12,79% e 10,99%, respectivamente. As componentes q_1 , q_2 , q_3 e q_4 da matriz **B** (que se refere aos ambientes) também explicam 48,49%, 21,63%, 12,79% e 10,99%, respectivamente, uma vez que a matriz **G** é diagonal. E a matriz **C**, que é formada por r_1 , explica 93,90% da $SQ_{G \times E \times A}$ (Tabela 4.14).

Tabela 4.15. Escores das componentes principais para o arranjo da interação tripla do modelo de Tucker3(4,4,1).

Arranjo Núcleo G				
	r_1			
	q_1	q_2	q_3	q_4
p_1	420285,3179	0,0000	0,0000	0,0000
p_2	0,0000	280708,3807	0,0000	0,0000
p_3	0,0000	0,0000	215854,8355	0,0000
p_4	0,0000	0,0000	0,0000	200119,9025

O arranjo núcleo **G**(4, 4, 1), apresentado na Tabela 4.15, mostra a relação entre as componentes e indica que a relação mais importante é entre as primeiras componentes de cada fator ($g_{111} = 420285, 3179$). Este valor indica que a combinação da primeira componente dos genótipos com a primeira componente dos ambientes e com a primeira componente dos atributos explicam 48,49% da $SQ_{G \times E \times A}$. A segunda relação mais importante é entre as segundas componentes dos genótipos e dos ambientes com a componente dos atributos ($g_{221} = 280708, 3807$), que explica 21,63% da soma de quadrados da interação tripla entre genótipos $\times$ ambientes $\times$ atributos. A terceira e quarta relações mais importantes são $g_{331} = 215854, 8355$ e $g_{441} = 200119, 9025$, que explicam 12,79% e 10,99%, respectivamente. As outras relações apresentaram valores iguais a zero, portanto não são úteis para explicar a variabilidade da interação tripla.

A componente r_1 da matriz **C** é caracterizada por um contraste entre AT4, e AT8. Assim, o *joint plot* da Figura 4.18 é projetado dentro da componente r_1 que explica 93,90% da $SQ_{G \times E \times A}$ e apresenta a interação entre genótipos e ambientes nos atributos AT4 e AT8. Assim, em relação ao atributo avaliado pelo método de escurecimento natural do grão no dia 90 (AT4), que apresentou valores negativos pela Tabela 4.14, observa-se pela Figura 4.18 que:

- o genótipo G17 teve uma interação negativa com o ambiente F (Campinas/águas) e positiva com o ambiente D (Ribeirão Preto/inverno);
- o genótipo G15 teve uma interação negativa com o ambiente D e positiva com o ambiente F;
- o genótipo G18 teve uma interação negativa com o ambiente C (Votuporanga/inverno) e o ambiente A (Campinas/seca);
- os genótipos G2 e G19 tiveram uma interação negativa com o ambiente E e positiva com o ambiente A e C.

Figura 4.18. *Joint plot* projetado dentro da primeira componente do terceiro modo para o modelo Tucker3(4,4,1).

Em relação em relação ao atributo avaliado pelo método de escurecimento acelerado do grão com 96 horas de exposição (AT8), que apresenta valores positivos pela Tabela 4.14, as interpretações são opostas as observadas anteriormente, isto é:

- o genótipo G17 teve uma interação positiva com o ambiente F (Campinas/águas) e negativa com o ambiente D (Ribeirão Preto/inverno);
- o genótipo G15 teve uma interação positiva com o ambiente D e negativa com o ambiente F;
- o genótipo G18 teve uma interação positiva com o ambiente C (Votuporanga/inverno) e o ambiente A (Campinas/seca);
- os genótipos G2 e G19 tiveram uma interação positiva com o ambiente E e negativa com o ambiente A e C.

Observa-se ainda que o genótipo G11, localizado na região central do *joint plot* apresentou desempenho médio nos atributos avaliados e que os ambientes C, D e F (setas longas) representaram mais a variabilidade da interação genótipos $\times$ ambientes associada ao componente do terceiro modo (atributos) do que os ambientes A, B e E (setas curtas).

Alguns resultados obtidos pelo modelo Tucker3(4,4,1) discordam daqueles observados nas análises AMMI dos atributos AT4 e AT8, uma vez que o genótipo G17 foi considerado estável para o AT8, porém o G15 foi o menos estável para o AT4, semelhante ao apresentado na Figura 4.18.

Observando a Figura 4.8, que apresenta o *biplot* para o AT4, nota-se que genótipo G15 interage positivamente com o ambiente F e negativamente com o ambiente D, da mesma forma que observado no *joint plot* da Figura 4.18. De forma semelhante, o G2 interage positivamente com o ambiente C e negativamente com o ambiente E em ambas as análises. Em relação aos genótipos mais estáveis, em ambas as análises aparecem os genótipos G3 e G11.

Para o AT8, analisando o *biplot* da Figura 4.16, o genótipo G2 interage positivamente com o ambiente C e negativamente com o ambiente E, sendo um comportamento oposto ao apresentado no *joint plot* da Figura 4.18. Mesma contradição ocorre para o G19 em relação ao ambiente A.

Considerando as adaptações específicas, para o AT4 em ambas as análises (AMMI individuais e Tucker3) observou-se uma interação positiva do genótipo G15 com o ambiente F e negativa com o ambiente D e uma interação positiva com o ambiente C e negativa com o ambiente E para o G19 e G2. Para o AT8, em ambas as análises observou-se uma interação positiva do G17 com o ambiente F e negativa com o ambiente D.

Estes resultados demonstram que o modelo Tucker3 foi eficiente para explorar a interação tripla entre genótipos $\times$ ambientes $\times$ atributos e identificar adaptações específicas e estabilidade genotípica, o que corrobora com os estudos conduzidos por Varela et al. (2006), Araújo et al. (2010), Araújo et al. (2011), Maia et al. (2014) e Garcia-Peña et al. (2016), que aplicaram a metodologia de Tucker3 para decompor a interação tripla.

4.3.2 Análise de Procrustes

A estatística de Procrustes foi calculada com a finalidade de comparar os *biplots* da análise AMMI individual dos atributos AT4 e AT8 com o *joint plot* do modelo Tucker3(4,4,1).

Tabela 4.16. Estatística de Procrustes (M^2) para o caso bivariado.

	Genótipos	Ambientes
<i>biplot</i> AT4 $\times$ <i>joint plot</i>	0,9675	0,0779
<i>biplot</i> AT8 $\times$ <i>joint plot</i>	0,9725	0,3613

Para os genótipos, o *joint plot* da análise Tucker3(4,4,1) e os *biplots* dos atributos AT4 e AT8 tiveram valores de M^2 similares após a transformação de Procrustes, porém para os ambientes, a similaridade das configurações foi maior para o AT4 (Tabela 4.16). Este resultado confirma as conclusões que foram consideradas na seção (4.3.1).

5 CONSIDERAÇÕES FINAIS

Os resultados obtidos pela aplicação da análise AMMI, modelo Tucker3 e análise de Procrustes permitem extrair as seguintes considerações:

- Nos estudos com mais de uma variável resposta, a análise AMMI feita individualmente para cada atributo não foi suficiente para evidenciar todas as relações existentes entre os genótipos e atributos;
- Quanto maior o número de atributos envolvidos no estudo, maior a dificuldade de uma interpretação global dos resultados das análises AMMI individuais, sendo necessário o uso de metodologias que analisem simultaneamente esses atributos.
- O método de Tucker3 é adequado para lidar com o arranjo triplo na análise da interação genótipos $\times$ ambientes $\times$ atributos e a interpretação de seus resultados apresentados no gráfico *joint plot* é imediata por ser semelhante à interpretação do gráfico *biplot* da análise AMMI;
- O gráfico *joint plot* permite a identificação de genótipos e ambientes estáveis e adaptações específicas, levando em consideração o caso multiatributo;
- Os resultados mostraram que é recomendável o uso do método de Tucker3 associado a análise de Procrustes quando se pretende estudar as similaridades entre as análises individuais e globais.

De um modo geral e avaliando os resultados finais deste trabalho, pode-se concluir que:

- O genótipo G15 (Gen 107-14A-336) apresentou uma adaptação específica ao ambiente F (Campinas/águas), os quais exibiram características semelhantes: alto valor médio de luminosidade e baixo valor de estabilidade;
- O genótipo G8 (Gen 45-2F-293P) apresentou em todas as análises individuais alto valor médio de luminosidade e pelo método de Tucker3 uma adaptação específica com o ambiente A (Campinas/seca), que apresentou alto valor médio de luminosidade em todos os atributos;
- O G12 (Gen 104-1A-291) exibiu alta estabilidade em todos os atributos e valor médio de luminosidade acima da média. Sua estabilidade também foi confirmada pela análise de Tucker3 no caso multiatributo, portanto este genótipo possui características de interesse no estudo da interação $G \times E \times A$.

REFERÊNCIAS

- Allard, R. W. e Bradshaw, A. D. (1964). Implications of Genotype-Environmental Interactions in Applied Plant Breeding. *Crop Science*, 4(5):503–508.
- Annicchiarico, P. (2002). *Genotype $\times$ environment interactions: challenges and opportunities for plant breeding and cultivar recommendations*. Rome: Food and Agriculture Organization.
- Arabie, P., Carroll, J., e DeSarbo, W. (1987). *Three-way scaling and clustering*. Sage Publications, Beverly Hills.
- Araújo, L. B., Nualles, M. V., Araujo, M. F. C., e Dias, C. T. S. (2010). Gráficos biplot e joint plot para o estudo da interação tripla. *Ciência Rural*, 40(4):833–839.
- Araújo, L. B., Nualles, M. V., Araujo, M. F. C., e Dias, C. T. S. (2011). Multiattribute response of maize genotypes tested in different coastal regions of brazil. *International Journal of Agronomy*, 2011:6.
- Araújo, L. C. A., Ramalho, M. A. P., e Abreu, A. F. B. (2012). Estimates of genetic parameters of late seed-coat darkening of carioca type dry beans. *Ciência e Agrotecnologia*, 36(2):156–162.
- Azevedo, L. B., Junior, V. C. A., e Fernandes, J. S. C. (2016). Transformação box-cox na homocedasticidade e normalidade uni e multivariada em experimentos de batata-doce. *Horticultura Brasileira*, 34(1):93–101.
- Basford, K. E., Kroonenberg, P. M., DeLacy, I. H., e Lawrence, P. K. (1990). Multiattribute evaluation of regional cotton variety trials. *Theoretical and Applied Genetics*, 79:225–234.
- Basford, K. E. e McLachlan, G. J. (1985). The mixture method of clustering applied to three-way data. *Journal of Classification*, 2:109–125.
- Box, G. E. P. e Cox, D. R. (1964). An analysis of transformations. *Journal of the Royal Statistical Society. Series B (Methodological)*, 26(2):211–252.
- Carroll, J. e Arabie, P. (1980). Multidimensional scaling. In Rosenzweig, M. e Porter, L., editors, *Annual Review of Psychology*, pages 607–649. Palto Alto.
- Chaves, L. J. (2001). Interação de genótipos com ambientes. In Nass, L. L., Valois, A. C. C., de Melo, I. S., e Valadares-Inglis, M. C., editors, *Recursos genéticos e melhoramento - plantas*, pages 673–713. Rondonópolis: Fundação MT.
- Cox, T. F. e Cox, M. A. A. (2000). *Multidimensional Scaling*. CRC Press, 2 edition.
- Crossa, J., Gauch, H. G., e Zobel, R. W. (1990). Additive main effects and multiplicative interaction analysis of two international maize cultivar trials. *Crop Science*, 30(3):493–500.
- Duarte, J. B. e Vencovsky, R. (1999). Interação genótipos $\times$ ambientes: uma introdução à análise AMMI. *Ribeirão Preto: Sociedade Brasileira de Genética*.
- Eberhart, S. A. e Russell, W. A. (1966). Stability parameters for comparing varieties. *Crop Science*, 6(1):36–40.
- Eckart, C. e Young, G. (1936). The approximation of one matrix by another of lower rank. *Psychometrika*, 1(3):211–218.
- Fan, X., Kang, M. S., Chen, H., Zhang, Y., Tan, J., e Xu, C. (2007). Yield stability of maize hybrids evaluated in multi-environment trials in yunnan, china. *Agronomy Journal*, 99:220–228.

- Fisher, R. A. (1918). The correlation between relatives on the supposition of mendalian inheritance. *Trans. R. Soc. Edinb.*, 52:399–433.
- Fisher, R. A. e Mackenzie, W. A. (1923). Studies in crop variation ii: The manurial response of different potato varieties. *Journal of Agricultural Science*, 23:311–320.
- Gabriel, K. R. (1971). The biplot graphic display of matrices with application to principal componente analysis. *Biometrika*, 58(3):453–467.
- Gabriel, K. R. (1978). Least squares approximation of matrices by additive and multiplicative models. *Journal of the Royal Statistical Society*, 40(2):186–196.
- Garcia-Peña, M., Arciniegas-Alarcon, S., Basford, K., e Dias, C. T. S. (2016). Analysis of sunflower data from a multi-attribute genotype $\times$ environment trial in brazil. *Communications in Biometry and Crop Science*, 11(2):127–139.
- Garcia-Peña, M. e Dias, C. T. S. (2009). Análise dos modelos aditivos com interação multiplicativa (AMMI) bivariados. *Rev. Bras. Biom*, 27(4):586–602.
- Gauch, H. G. (1992). *Statistical Analysis of Regional Yield Trials: AMMI Analysis of Factorial Designs*. Elsevier.
- Gauch, H. G. (2013). A simple protocol for AMMI analysis of yield trials. *Crop Science*, 53(5):1860–1869.
- Gauch, H. G. e Zobel, R. W. (1996). Ammi analysis of yield trials. In Kang, M. S. e Gauch, H. G., editors, *Genotype-by-Environment Interaction*, pages 85–122. Florida: CRC Press.
- Gollob, H. F. (1968). A statistical model which combines features of factor analytic and analysis of variance techniques. *Psychometrika*, 33:73–115.
- Henderson, C. R. (1975). Best linear estimation and prediction under a selection model. *Biometrics*, 31(2):423–447.
- Kang, M. S. (1993). Issues in GE interaction. In Rao, V., Hanson, I. E., e Rajanaidu, N., editors, *Genotype–Environment Interaction Studies in Perennial Tree Crops*, pages 67–73. Palm Oil Research Institute of Malaysia.
- Kang, M. S. (2001). Genotype–environment interaction: Progress and prospects. In Kang, M. S., editor, *Quantitative Genetics, Genomics and Plant Breeding*, pages 221–244. CABI Publishing.
- Kempton, R. A. (1984). The use of biplots in interpreting variety by environment interactions. *Journal of Agricultural Science*, 103:123–135.
- Koundinya, A. V., Pandit, M. K., Ramesh, D., e Mishra, P. (2019). Phenotypic stability of eggplant for yield and quality through AMMI, GGE and cluster analyses. *Scientia Horticulturae*, 247:216–223.
- Kroonenberg, P. M. (2008). *Applied Multiway Data Analysis*. John Wiley & Sons.
- Kroonenberg, P. M. e Basford, K. E. (1989). An investigation of multi-attribute genotype response across environments using three-mode principal component analysis. *Euphytica*, 44(1-2):109–123.
- Kroonenberg, P. M. e Leeuw, J. (1980). Principal component analysis of three-mode data by means of alternating least squares algorithms. *Psychometrika*, 45(1):69–97.
- Kruskal, J. (1977). The relationship between multidimensional scaling and clustering. In Ryzin, J. V., editor, *Classification and Clustering*, pages 17–44. Academic Press, New York.

- Krzanowski, W. J. (2000). *Principles of multivariate analysis: a user's perspective*. Oxford University Press.
- Levin, J. (1965). Three-mode factor analysis. *Psychological Bulletin*, 64:442–452.
- Lohmoller, J.-B. (1979). Tucker's three-mode factor analysis: Scaling, rotations, other models. *Archiv fur die Psychologie*, 131:137–166.
- López-Morales, F., Vázquez-Carrillo, M. G., Molina-Galán, J. D., García-Zavala, J. J., Corona-Torres, T., Cruz-Izquierdo, S., López-Romero, G., Reyes-López, D., e Esquivel-Esquivel, G. (2017). Interacción genotipo-ambiente, estabilidad del rendimiento y calidad de grano en maíz Tuxpeño. *Revista Mexicana de Ciencias Agrícolas*, 8(5):1035–1050.
- Maia, M. C. C., Vello, N. A., de Araújo, L. B., dos Santos Dias, C. T., de Oliveira, L. C., e de Moura Rocha, M. (2014). Interação genótipo-ambiente em soja via análise de componentes principais com múltiplas matrizes de dados. *Revista Agro@mbiente*, 8(1):104–111.
- Mandel, J. A. (1971). A new analysis of variance model for non-additive data. *Technometrics*, 13:1–18.
- Matkovic Stojšin, M., Zecevic, V., Knezevic, D., Boskovic, J., e Micanovic, D. (2018). Variability, AMMI and cluster analysis for quality components of different wheat genotypes. *AGROFOR International Journal*, 3(2):65–72.
- Neisse, A. C. e Hongyu, K. (2016). Aplicação de Componentes Principais e Análise Fatorial a Dados Criminais de 26 Estados dos Eua. *Engineering and Science*, 2(6):105–115.
- Ngailo, S., Shimelis, H., Sibiya, J., Mtunda, K., e Mashilo, J. (2019). Genotype-by-environment interaction of newly-developed sweet potato genotypes for storage root yield, yield-related traits and resistance to sweet potato virus disease. *Heliyon*, 5(3):1–23.
- Onofri, A. e Ciricifolo, E. (2007). Using R to perform the AMMI analysis on agriculture variety trials. *R News*, 7(1):14–19.
- Pearson, K. (1901). On lines and planes of closest fit to systems of points in space. *Philosophical Magazine*, 2(11):559–572.
- Perkins, J. M. e Jinks, J. L. (1968). Environmental and genotype-environmental components of variability. iii. multiple lines and crosses. *Heredity*, 23:339–356.
- Pike, E. W. e Silverberg, T. R. (1952). Designing mechanical computers. *Machine Design*, 24.
- R Core Team (2019). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.
- Rencher, A. C. e Schaalje, G. B. (2007). *Linear Models in Statistics*. John Wiley & Sons, 2 edition.
- Resende, M. D. V. (2007). *SELEGENREML/BLUP: sistema estatístico e seleção genética computadorizada via modelos lineares mistos*. Colombo: Embrapa Florestas.
- Ribeiro, N. D., da Silva Domingues, L., Gruhn, E. M., Zemolin, A. E. M., e de Abreu Rodrigues, J. (2014). Desempenho agrônomo e qualidade de cozimento de linhagens de feijão de grãos especiais. *Revista Ciência Agronômica*, 45(1):92–100.
- Ribeiro, N. D., Jost, E., e Filho, A. C. (2004). Efeitos da interação genótipo $\times$ ambiente no ciclo e na coloração do tegumento dos grãos do feijoeiro comum. *Bragantia*, 63(3):373–380.

- Rodrigues, P. C., Malosetti, M., Gauch, H. G., e van Eeuwijk, F. A. (2014). A weighted AMMI algorithm to study genotype-by-environment interaction and QTL-by-environment interaction. *Crop Science*, 54(4):1555–1570.
- Schepers, J., Ceulemans, E., e Mechelen, I. (2008). Selecting among multi-mode partitioning models of different complexities: A comparison of four model selection criteria. *Journal of Classification*, 25(1):67–85.
- Silva, G. O., Carvalho, A. D. F., Vieira, J. V., e Benin, G. (2011). Verificação da adaptabilidade e estabilidade de populações de cenoura pelos métodos AMMI, GGE biplot e REML/BLUP. *Bragantia*, 70(3):494–501.
- Smilde, A., Bro, R., e Geladi, P. (2004). *Multi-way Analysis with Applications in the Chemical Sciences*. John Wiley & Sons.
- Smith, A. B., Cullis, B. R., e Thompson, R. (2005). The analysis of crop cultivar breeding and evaluation trials: an overview of current mixed model approaches. *Journal of Agricultural Science*, 143(6):449–462.
- Souza, J. (1988). Análise de componentes principais e suas aplicações: métodos estatísticos nas ciências psicossociais. *Thesaurus*, 2:68.
- Spitti, A. M. D. S., Carbonell, S. A. M., dos Santos Dias, C. T., Sabino, L. G., Carvalho, C. R. L., e Chiorato, A. F. (2019). Carioca bean genotypes for tolerance to grain darkening by natural and accelerated methods. *Ciência e Agrotecnologia*, 43:1–10.
- Timmerman, M. E. e Kiers, H. A. L. (2000). Three-mode principal components analysis: Choosing the numbers of components and sensitivity to local optima. *British Journal of Mathematical and Statistical Psychology*, 53(1):1–16.
- Tucker, L. R. (1964). The extension of factor analysis to three-dimensional matrices. In Frederiksen N, G. H., editor, *Contributions to Mathematical Psychology*, pages 110–182. Holt, Rinehart & Winston, New York.
- Tucker, L. R. (1966). Some mathematical notes on three-mode factor analysis. *Psychometrika*, 31.
- Varela, M., Crossa, J., Rane, J., Joshi, A. K., e Trethowan, R. (2006). Analysis of a three-way interaction including multi-attributes. *Australian Journal of Agricultural Research*, 57:1185–1193.
- Varela, M. e Torres, V. (2005). Aplicación del análisis de componentes principales de tres modos en la caracterización multivariada de somaclones de king-grass. *Revista Cubana de Ciencia Agrícola*, 39:543–552.
- Wade, L., McLaren, C., Quintana, L., Harnpichitvitaya, D., Rajatasereekul, S., Sarawgi, A., Kumar, A., Ahmed, H., Sarwoto, Singh, A., Rodriguez, R., Siopongco, J., e Sarkarung, S. (1999). Genotype by environment interactions across diverse rainfed lowland rice environments. *Field Crops Research*, 64:35–50.
- Williams, E. J. (1952). The interpretation of interactions in factorial experiments. *Biometrika*, 39:65–81.
- Yan, W. e Holland, J. B. (2010). A heritability-adjusted GGE biplot for test environment evaluation. *Euphytica*, 171:355–369.
- Yan, W. e Kang, M. S. (2003). *GGE Biplot Analysis: A Graphical Tool for Breeders, Geneticists and Agronomists*. CRC Press.

Zobel, R. W., Wright, M. J., e Gauch, H. G. (1988). Statistical analysis of a yield trial. *Agronomy Journal*, 80:388–393.

ANEXOS

Anexo A

As operações de translação, rotação e dilatação da análise de Procrustes buscam minimizar a soma de quadrados das distâncias entre os pontos, que é dada por:

$$M^2 = \sum_{r=1}^n (\mathbf{y}_r - \rho \mathbf{A}^T \mathbf{x}_r - \mathbf{b})^T (\mathbf{y}_r - \rho \mathbf{A}^T \mathbf{x}_r - \mathbf{b})$$

5.1 Correspondência por translação

Sejam os pontos $\mathbf{x}_0$ e $\mathbf{y}_0$ os centróides das duas configurações, isto é

$$\mathbf{x}_0 = \frac{1}{n} \sum_{r=1}^n \mathbf{x}_r \quad \mathbf{y}_0 = \frac{1}{n} \sum_{r=1}^n \mathbf{y}_r$$

Adicionando e subtraindo $\mathbf{x}_0$ e $\mathbf{y}_0$ em (2.23), tem-se:

$$\begin{aligned} M^2 &= \sum_{r=1}^n \left((\mathbf{y}_r - \mathbf{y}_0) - \rho \mathbf{A}^T (\mathbf{x}_r - \mathbf{x}_0) + \mathbf{y}_0 - \rho \mathbf{A}^T \mathbf{x}_0 - \mathbf{b} \right)^T \left((\mathbf{y}_r - \mathbf{y}_0) - \rho \mathbf{A}^T (\mathbf{x}_r - \mathbf{x}_0) + \mathbf{y}_0 - \rho \mathbf{A}^T \mathbf{x}_0 - \mathbf{b} \right) \\ M^2 &= \sum_{r=1}^n \left((\mathbf{y}_r - \mathbf{y}_0) - \rho \mathbf{A}^T (\mathbf{x}_r - \mathbf{x}_0) \right)^T \left((\mathbf{y}_r - \mathbf{y}_0) - \rho \mathbf{A}^T (\mathbf{x}_r - \mathbf{x}_0) \right) \\ &\quad + n \left(\mathbf{y}_0 - \rho \mathbf{A}^T \mathbf{x}_0 - \mathbf{b} \right)^T \left(\mathbf{y}_0 - \rho \mathbf{A}^T \mathbf{x}_0 - \mathbf{b} \right) \end{aligned} \quad (5.1)$$

Como o último termo da equação (5.1) é não negativo e $\mathbf{b}$ ocorre somente neste termo, para que M^2 seja mínimo, $\mathbf{b} = \mathbf{y}_0 - \rho \mathbf{A}^T \mathbf{x}_0$, consequentemente:

$$\mathbf{x}_r^* = \rho \mathbf{A}^T (\mathbf{x}_r - \mathbf{x}_0) + \mathbf{y}_0$$

o que implica que o centróide no espaço $\mathbf{X}^*$ é coincidente com o centróide no espaço $\mathbf{Y}$. A maneira mais conveniente de garantir isso é inicialmente translacionar as configurações no espaço $\mathbf{X}$ e no espaço $\mathbf{Y}$ para que ambos tenham seus centróides na origem.

5.2 Correspondência por dilatação

Assumindo $\mathbf{x}_0 = \mathbf{y}_0 = 0$, então:

$$\begin{aligned} M^2 &= \sum_{r=1}^n (\mathbf{y}_r - \rho \mathbf{A}^T \mathbf{x}_r)^T (\mathbf{y}_r - \rho \mathbf{A}^T \mathbf{x}_r) \\ &= \sum_{r=1}^n \mathbf{y}_r^T \mathbf{y}_r + \rho^2 \sum_{r=1}^n \mathbf{x}_r^T \mathbf{x}_r - 2\rho \sum_{r=1}^n \mathbf{x}_r^T \mathbf{A} \mathbf{y}_r \\ &= tr(\mathbf{Y} \mathbf{Y}^T) + \rho^2 tr(\mathbf{X} \mathbf{X}^T) - 2\rho tr(\mathbf{X} \mathbf{A} \mathbf{Y}^T) \end{aligned} \quad (5.2)$$

diferenciando em relação a ρ dado $\hat{\rho}$ e igualando a zero, o valor de ρ que minimiza M^2 é:

$$\hat{\rho} = tr(\mathbf{X} \mathbf{A} \mathbf{Y}^T) / tr(\mathbf{X} \mathbf{X}^T) = tr(\mathbf{A} \mathbf{Y}^T \mathbf{X}) / tr(\mathbf{X} \mathbf{X}^T) \quad (5.3)$$

A matriz de rotação $\mathbf{A}$ ainda é desconhecida e será considerada a seguir.

5.3 Correspondência por rotação

O valor de M^2 na equação (5.2) será um mínimo se $tr(\mathbf{X}\mathbf{A}\mathbf{Y}^T) = tr(\mathbf{A}\mathbf{Y}^T\mathbf{X})$ for um máximo. Seja $\mathbf{C} = \mathbf{Y}^T\mathbf{X}$ e considere que $\mathbf{C}$ é igual a decomposição por valores singulares, isto é, $\mathbf{C} = \mathbf{U}\mathbf{\Lambda}\mathbf{V}^T$, em que $\mathbf{U}$ e $\mathbf{V}$ são as matrizes ortonormais e $\mathbf{\Lambda}$ a matriz diagonal com os valores singulares, então:

$$tr(\mathbf{A}\mathbf{C}) = tr(\mathbf{A}\mathbf{U}\mathbf{\Lambda}\mathbf{V}^T) = tr(\mathbf{V}^T\mathbf{A}\mathbf{U}\mathbf{\Lambda})$$

Logo $\mathbf{A}$, $\mathbf{V}$ e $\mathbf{U}$ são matrizes ortonormais e, conseqüentemente, $\mathbf{V}^T\mathbf{A}\mathbf{U}$ também é ortonormal. Uma vez que $\mathbf{\Lambda}$ é diagonal e uma matriz ortonormal não pode ter qualquer elemento maior que a unidade, então:

$$tr(\mathbf{A}\mathbf{C}) = tr(\mathbf{V}^T\mathbf{A}\mathbf{U}\mathbf{\Lambda}) \leq tr(\mathbf{\Lambda})$$

Portanto, M^2 é minimizado quando $tr(\mathbf{A}\mathbf{C}) = tr(\mathbf{\Lambda})$, implicando que:

$$\mathbf{V}^T\mathbf{A}\mathbf{U}\mathbf{\Lambda} = \mathbf{\Lambda} \quad (5.4)$$

A equação (5.4) tem solução $\mathbf{A} = \mathbf{V}\mathbf{U}^T$ resultando a matriz de rotação $\mathbf{A}$ como o produto das matrizes ortonormais da decomposição por valores singulares de $\mathbf{Y}^T\mathbf{X}$.

Pré-multiplicando (5.4) por $\mathbf{V}$ e pós-multiplicando por $\mathbf{V}^T$, têm-se $\mathbf{A}\mathbf{U}\mathbf{\Lambda}\mathbf{V}^T = \mathbf{V}\mathbf{\Lambda}\mathbf{V}^T$, conseqüentemente:

$$\mathbf{A}\mathbf{C} = \mathbf{A}\mathbf{U}\mathbf{\Lambda}\mathbf{V}^T = \mathbf{V}\mathbf{\Lambda}\mathbf{V}^T = (\mathbf{V}\mathbf{\Lambda}^2\mathbf{V}^T)^{1/2}$$

e como $\mathbf{U}\mathbf{U}^T = \mathbf{U}^T\mathbf{U} = \mathbf{I}$, sendo $\mathbf{I}$ uma matriz identidade, então:

$$\mathbf{A}\mathbf{C} = (\mathbf{V}\mathbf{\Lambda}^2\mathbf{V}^T)^{1/2} = (\mathbf{V}\mathbf{\Lambda}\mathbf{U}\mathbf{U}^T\mathbf{V}^T)^{1/2} = (\mathbf{C}^T\mathbf{C})^{1/2} \quad (5.5)$$

Logo, a matriz de rotação ótima $\mathbf{A}$ é dada por

$$\mathbf{A} = (\mathbf{C}^T\mathbf{C})^{1/2}\mathbf{C}^{-1} = (\mathbf{X}^T\mathbf{Y}\mathbf{Y}^T\mathbf{X})^{1/2}(\mathbf{Y}^T\mathbf{X})^{-1}$$

Como $\mathbf{C} = \mathbf{Y}^T\mathbf{X}$ e por (5.3), o parâmetro de dilatação será dado por:

$$\hat{\rho} = tr(\mathbf{X}^T\mathbf{Y}\mathbf{Y}^T\mathbf{X})^{1/2} / tr(\mathbf{X}^T\mathbf{X})$$

Sabendo que $tr(\mathbf{A}\mathbf{B}) = tr(\mathbf{B}\mathbf{A})$ e $tr(\mathbf{A}\mathbf{A}^T) = tr(\mathbf{A}^T\mathbf{A})$, a avaliação da correspondência das duas configurações pode ser feita aplicando o valor minimizado de $\hat{\rho}$ em (5.2), isto é

$$M^2 = tr(\mathbf{Y}\mathbf{Y}^T) + \rho^2 tr(\mathbf{X}\mathbf{X}^T) - 2\rho tr(\mathbf{X}\mathbf{A}\mathbf{Y}^T)$$

$$M^2 = tr(\mathbf{Y}\mathbf{Y}^T) + \left[\frac{tr(\mathbf{X}^T\mathbf{Y}\mathbf{Y}^T\mathbf{X})^{1/2}}{tr(\mathbf{X}^T\mathbf{X})} \right]^2 tr(\mathbf{X}\mathbf{X}^T) - 2 \left[\frac{tr(\mathbf{X}^T\mathbf{Y}\mathbf{Y}^T\mathbf{X})^{1/2}}{tr(\mathbf{X}^T\mathbf{X})} \right] tr(\mathbf{X}\mathbf{A}\mathbf{Y}^T)$$

$$M^2 = tr(\mathbf{Y}\mathbf{Y}^T) + \frac{[tr(\mathbf{X}^T\mathbf{Y}\mathbf{Y}^T\mathbf{X})^{1/2}]^2}{tr(\mathbf{X}^T\mathbf{X})} - 2 \frac{tr(\mathbf{X}^T\mathbf{Y}\mathbf{Y}^T\mathbf{X})^{1/2}}{tr(\mathbf{X}^T\mathbf{X})} tr(\mathbf{X}\mathbf{A}\mathbf{Y}^T)$$

Sabendo também que $tr(\mathbf{A}\mathbf{B}\mathbf{C}) = tr(\mathbf{B}\mathbf{C}\mathbf{A}) = tr(\mathbf{C}\mathbf{A}\mathbf{B})$, que $\mathbf{A}\mathbf{C} = \mathbf{A}\mathbf{Y}^T\mathbf{X}$, que $(\mathbf{A}\mathbf{B})^T = \mathbf{B}^T\mathbf{A}^T$ e pela equação (5.5), tem-se que:

$$\text{tr}(\mathbf{X}\mathbf{A}\mathbf{Y}^T) = \text{tr}(\mathbf{A}\mathbf{Y}^T\mathbf{X}) = \text{tr}(\mathbf{A}\mathbf{C}) = \text{tr}(\mathbf{C}^T\mathbf{C})^{1/2} = \text{tr}(\mathbf{X}^T\mathbf{Y}\mathbf{Y}^T\mathbf{X})^{1/2}$$

$$M^2 = \text{tr}(\mathbf{Y}\mathbf{Y}^T) + \frac{\left[\text{tr}(\mathbf{X}^T\mathbf{Y}\mathbf{Y}^T\mathbf{X})^{1/2}\right]^2}{\text{tr}(\mathbf{X}^T\mathbf{X})} - 2\frac{\text{tr}(\mathbf{X}^T\mathbf{Y}\mathbf{Y}^T\mathbf{X})^{1/2}}{\text{tr}(\mathbf{X}^T\mathbf{X})}\text{tr}(\mathbf{X}^T\mathbf{Y}\mathbf{Y}^T\mathbf{X})^{1/2}$$

$$M^2 = \text{tr}(\mathbf{Y}\mathbf{Y}^T) - \frac{\{\text{tr}(\mathbf{X}^T\mathbf{Y}\mathbf{Y}^T\mathbf{X})^{1/2}\}^2}{\text{tr}(\mathbf{X}^T\mathbf{X})}$$

Agora é possível modificar a escala dividindo M^2 por $\text{tr}(\mathbf{Y}^T\mathbf{Y})$ e, assim, obtém-se estatística de Procrustes:

$$M^2 = 1 - \frac{\{\text{tr}(\mathbf{X}^T\mathbf{Y}\mathbf{Y}^T\mathbf{X})^{1/2}\}^2}{\text{tr}(\mathbf{X}^T\mathbf{X})\text{tr}(\mathbf{Y}^T\mathbf{Y})}$$

que indica maior similaridade entre as configurações, quanto menor for o seu valor.