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

Métodos de correção de autovalores e regressão isotônica nos modelos
AMMI

Lúcio Borges de Araújo

Dissertação apresentada para obtenção do título de Mestre
em Agronomia. Área de concentração: Estatística e Experi-
mentação Agronômica

Piracicaba
2005

Lúcio Borges de Araújo
Licenciado em Matemática

**Métodos de correção de autovalores e regressão isotônica nos modelos
AMMI**

Orientador:
Prof. Dr. **CARLOS TADEU DOS SANTOS DIAS**

Dissertação apresentada para obtenção do título de Mestre
em Agronomia. Área de concentração: Estatística e Experi-
mentação Agronômica

**Piracicaba
2005**

Dados Internacionais de Catalogação na Publicação (CIP)
DIVISÃO DE BIBLIOTECA E DOCUMENTAÇÃO - ESALQ/USP

Araújo, Lúcio Borges de
Métodos de correção de autovalores e regressão isotônica nos modelos AMMI / Lúcio
Borges de Araújo. - - Piracicaba, 2005.
75 p.

Dissertação (Mestrado) - - Escola Superior de Agricultura Luiz de Queiroz, 2005.

1. Análise de regressão e de correlação 2. Análise de variância 3. Estabilidade
(teoria de sistemas e controle) 4. Interação genótipo – Ambiente I. Título

CDD 519.5

“Permitida a cópia total ou parcial deste documento, desde que citada a fonte – O autor”

DEDICATÓRIA

A **Deus** pois,

“Em seu coração o homem planeja o seu caminho,

*mas o **Senhor** determina os seus passos”*(Provérbios 16:9) e

“que adianta ao homem ganhar o mundo inteiro e perder sua alma?”(Marcos 8:36)

Aos meus pais **Luis Guilhermme de Araújo** e **Tânia Maria Borges Araújo**,
que sempre lutaram para dar uma boa educação e formação aos seus filhos

Aos meus irmãos **Gabriel**, **Aurélia** e **Evaldo**
pela constante amizade e grandes incentivos

A namorada **Mírian**, que amo muito
pelo incentivo e apoio constante nas horas mais difíceis.

Agradecimentos

A ESALQ/USP pela estrutura física e humana disponível.

Ao professor Dr. Carlos Tadeu dos Santos Dias, pela amizade, exemplo de vida e principalmente pela orientação e estímulos à elaboração deste trabalho.

A Capes pela concessão de bolsas de estudos.

Aos membros da Igreja Evangélica na Paulista que me acolheram muito bem e são a minha família em Piracicaba.

Ao conselho do programa de Pós-graduação em Estatística e Experimentação Agrônômica, os professores Dra. Clarice Garcia Borges Demétrio, Dr. Décio Barbin e Dra. Roseli Aparecida Leandro, pelas valiosas sugestões e confiança.

Aos professores do Departamento de Ciências Exatas da ESALQ/USP, Dr. César Gonçalves de Lima, Dr. Edwin Moisés Ortega, Dr. Gabriel Adrián Sarriés, Dra. Maria Cristina Stolf Nogueira, Dr. Silvio Sandoval Zocchi e Dra. Sônia Maria Stefano Piedade, pela amizade e formação.

Aos funcionários do Departamento de Ciências Exatas da ESALQ/USP, as secretárias Solange de Assis Paes Sabadin e Luciane Brajão e o técnico em informática Jorge Alexandre Wiendl, pelos auxílios permanentes.

Aos colegas de estudo do mestrado Alexandre, Ana Paula, Angela, Cristiane, Édila, Elizabeth Toledo, Fernanda, Francine, Hélio, Joseane, Juliana Betini, Júlio, Melissa, Moita, Pâmela e Sandra e aos colegas do doutorado Adriano, Afrânio, Ana Maria, Andréia, Antônio Carlos, Cesar, David, Denise, Elizabeth Strapasson, Genevile, Giovana, Idemauro, João Maurício, Juliana Cespedes, Luciana, Luciano, Milton, Renato e Telde.

Um agradecimento especial reservo ao amigo Osmar Jesus Macedo, pelo convívio na mesma casa, exemplo de vida e amizade.

A todas as pessoas que contribuíram direta ou indiretamente para a realização deste trabalho.

SUMÁRIO

	Página
RESUMO	7
ABSTRACT	8
LISTA DE TABELAS	9
1 INTRODUÇÃO	11
2 DESENVOLVIMENTO	13
2.1 Revisão de Literatura	13
2.1.1 Interação de genótipos $\times$ ambientes	13
2.1.2 Modelos AMMI	16
2.1.3 Autovalores e Autovetores de uma matriz	19
2.1.3.1 Estimação de autovalores e distribuição amostral	20
2.1.3.2 Viés e Correção de autovalores	21
2.1.4 Regressão Isotônica	21
2.1.4.1 Regressão Isotônica: Um caso simples de dados ordenados	22
2.1.4.2 O algoritmo PAVA	23
2.1.4.3 Regressão isotônica em um conjunto quase-ordenado	24
2.1.4.4 Regressão isotônica generalizada	24
2.1.5 Decomposição em valores singulares	25
2.2 Material e Métodos	27
2.2.1 Características dos dados	27
2.2.2 Análises de variância	27
2.2.2.1 Análise de variância em cada ambiente	27
2.2.2.2 Análise de variância conjunta	29
2.2.3 Análises AMMI	31
2.2.4 Correções dos autovalores em matrizes $(GE)(GE)^t$ ou $(GE)^t(GE)$	33
2.2.5 Algoritmo para Isotonização de Stein	35

	6
2.2.5.1	Número de repetições 36
2.3	Resultados e Discussão 38
2.3.1	Experimento 1 38
2.3.1.1	Análise de variância conjunta 38
2.3.1.2	Análise AMMI sem correção dos autovalores 38
2.3.1.3	Análise AMMI com correção dos autovalores 39
2.3.1.4	Número de Repetições 42
2.3.2	Experimento 2 43
2.3.2.1	Análise de variância conjunta 43
2.3.2.2	Análise AMMI sem correção dos autovalores 44
2.3.2.3	Análise AMMI com correção dos autovalores 44
2.3.2.4	Número de Repetições 47
2.3.3	Experimento 3 48
2.3.3.1	Análise de variância conjunta 48
2.3.3.2	Análise AMMI sem correção dos autovalores 48
2.3.3.3	Análise AMMI com correção dos autovalores 49
2.3.3.4	Número de Repetições 51
2.3.4	Experimento 4 52
2.3.4.1	Análise de variância conjunta 52
2.3.4.2	Análise AMMI sem correção dos autovalores 52
2.3.4.3	Análise AMMI com correção dos autovalores 53
2.3.4.4	Número de Repetições 55
2.3.5	Considerações gerais 55
3	CONCLUSÕES 58
	REFERÊNCIAS 59
	ANEXOS 65

RESUMO

Métodos de correção de autovalores e regressão isotônica nos modelos AMMI

Em experimentação agrícola, é freqüente a necessidade de análise conjunta de grupos de experimentos. Em muitos casos, o pesquisador deseja generalizar resultados para condições gerais de regiões e/ou em avaliar o desempenho de vários genótipos (tratamentos) em diversos ambientes (locais e/ou ano). Quando um conjunto de experimentos é planejado para vários locais é necessário considerar o delineamento individual em cada local e a combinação total dos genótipos com os locais (interação genótipo $\times$ ambiente). Logo, os dados observados podem ser organizados em uma tabela de dupla entrada. Existem várias metodologias de análise e interpretação para a interação genótipo $\times$ ambiente proveniente de um grupo de cultivares testados em vários ambientes. Entre essas metodologias destaca-se os modelos AMMI (“additive main effects and multiplicative interaction model”), como o próprio nome diz é um método uni-multivariado, que engloba uma análise de variância para os efeitos principais, que são os efeitos dos genótipos (G) e os ambientes (E) e para efeitos multiplicativos (interação genótipo $\times$ ambiente), utiliza-se a decomposição em valor singular (DVS). Essa técnica multivariada baseia-se no uso dos autovalores e autovetores provenientes da matriz de interação genótipo $\times$ ambiente. Araújo e Dias (2005) verificaram o problema de superestimação e subestimação de autovalores estimados da maneira convencional. Para superar esses problemas de estimação de autovalores, Muirhead (1987) apresenta três métodos para corrigir autovalores estimados a partir das matrizes de covariâncias amostral e alerta que nem sempre essas correções mantêm a ordem decrescente de valores, assim é sugerido o uso de regressão isotônica para ordenar esses dados, mas propriamente um algoritmo apresentado por Lin e Pearlman (1985). Os resultados indicaram que: A regressão isotônica juntamente com o algoritmo foi necessária e se mostrou muito importante em todos conjuntos de dados; Houve uma redução no número de componentes significativos para serem retidos nos modelos, fazendo com que os modelos AMMI selecionados sejam mais parcimoniosos quando se utiliza qualquer um dos métodos de correção; O método 2 apresentou as menores taxa de correção da soma de quadrados da interação genótipo $\times$ ambiente e o método 3 apresentou a maiores taxa de correção; Em relação a medida $RMSPD_{PRESS}$, os menores valores foram obtidos quando se utilizou o método de correção 2. Já o método de correção 3 apresentou os maiores valores para $RMSPD_{PRESS}$; O método 2 também se mostrou melhor quando o interesse era verificar o ganho em número de repetições, sendo que este benefício esteve sempre próximo de 3 repetições. Já o método 3 é o que apresenta um menor ganho em número de repetições, em torno de duas repetições.

Palavras-chaves: Correção de autovalores; regressão isotônica; modelos AMMI; interação genótipo $\times$ ambiente; eficiência da correção;

ABSTRACT

Methods of eigenvalue correction and isotonic regression in models AMMI

In agricultural research is common to analyse groups of experiments. In many cases, the researcher intends to generalize results to general conditions of areas and/or evaluate the responses of several genotypes (treatments) in several environments (places and/or years). When a group of experiments is planned for several places it is necessary to consider the of design in each place and the combinations of the genotypes with the places (the interaction of genotype $\times$ environment). The observed data can be organized in an array. There are several methods of analysis and interpretation for the genotype $\times$ environment interaction from a group of genotype tested in several environments. These methods include AMMI models (“additive main effect and multiplicative interaction models”). As the name says it is a uni-multivariate method, that includes an analysis of variance for the main effects (the effects of the genotypes (G) and environments (E)) and assumes multiplicative effects for the genotype $\times$ environment interaction, using a singular value decomposition (DVS). This method estimates the eigenvalues and eigenvectors deriving from the matrix of genotype $\times$ environment interaction. Araújo and Dias (2005) found an overestimation and underestimation problem with the eigenvalues in the conventional way. To correct these problems Muirhead (1987) presents three methods to correct the eigenvalues from covariance the matrix and noted that these do not always maintain the order of values. The author suggested the use of isotonic regression to correct the eigenvalues, using an algorithm presented by Lin and Pearlman (1985). The results indicated that: The isotonic regression with the algorithm is necessary and it showed very important in all groups of data; There was a reduction in the number of significant components to be kept in the models and the order that the AMMI model selected is more parsimonious when any of the correction methods is used; The method 2 has the smallest rate of correction to the sum of squares of the genotype $\times$ environment interaction and method 3 has the largest correction rate; The measure $RMSPD_{PRESS}$ was smallest when method of correction 2 was used. The method of correction 3 has the largest values for $RMSPD_{PRESS}$; Method 2 was also better when the interest was to verify the gain in number of replicates, and this benefit was always close to 3 replicates. The method 3 gives the smaller gain in the number of replicates, of around two replicates.

Keywords: Correction of eigenvalues; isotonic regression; AMMI models; genotype $\times$ environment interaction; efficiency replicates;

LISTA DE TABELAS

		Página
1	Esquema da análise de variância para experimento de um grupo de genótipos avaliados em um local em r blocos	28
2	Esquema da análise de variância para experimento de um grupo de genótipos avaliados em um local em r repetições	29
3	Esquema da análise de variância para experimentos de um mesmo grupo de genótipos avaliado em e locais com r blocos	30
4	Esquema da análise de variância para experimentos de um mesmo grupo de genótipos avaliado em e locais com r repetições	31
5	Análise de variância conjunta do Experimento 1 com 20 genótipos avaliado em 34 ambientes com 4 blocos	38
6	Decomposição de soma de quadrados de interação do Experimento 1 com 20 genótipos avaliado em 34 ambientes com 4 blocos	39
7	Correção dos autovalores da matriz $(GE)(GE)^t$ do experimento 1 e os autovalores ajustados pela regressão isotônica	40
8	Análise do Teste F para os componentes corrigidos do Experimento 1 pelo método 1	41
9	Análise do Teste F para os componentes corrigidos do Experimento 1 pelo método 2	42
10	Análise do Teste F para os componentes corrigidos do Experimento 1 pelo método 3	42
11	$RMSPD_{PRESS}$ e R_{AMMI} para o melhor modelo AMMI selecionado para os dados do Experimento 1 após a correção dos autovalores pelos métodos 1, 2 e 3	43
12	Análise de variância conjunta do Experimento 2 com 16 genótipos avaliado em 24 ambientes com 4 blocos	43
13	Decomposição de soma de quadrados de interação do Experimento 2 com 16 genótipos avaliados em 24 ambientes com 4 blocos	44
14	Correção dos autovalores da matriz $(GE)(GE)^t$ do experimento 2 e os autovalores ajustados pela regressão isotônica	45
15	Análise do Teste F para os componentes do Experimento 2 corrigidos pelo método 1	46
16	Análise do Teste F para os componentes do Experimento 2 corrigidos pelo método 2	46
17	Análise do Teste F para os componentes do Experimento 2 corrigidos pelo método 3	47
18	$RMSPD_{PRESS}$ e R_{AMMI} para o melhor modelo AMMI selecionado para os dados do Experimento 2 após a correção dos autovalores pelos métodos 1, 2 e 3	47

19	Análise de variância conjunta do Experimento 3 com 9 genótipos avaliado em 20 ambientes com 4 blocos	48
20	Decomposição de soma de quadrados de interação do Experimento 3 com 9 genótipos avaliado em 20 ambientes com 4 blocos	49
21	Correção dos autovalores da matriz $(GE)(GE)^t$ do experimento 3 e os autovalores ajustados pela regressão isotônica	49
22	Análise do Teste F para os componentes do Experimento 3 corrigidas pelo método 1	50
23	Análise do Teste F para os componentes do Experimento 3 corrigidas pelo método 2	51
24	Análise do Teste F para os componentes do Experimento 3 corrigidas pelo método 3	51
25	$RMSPD_{PRESS}$ e R_{AMMI} para o melhor modelo AMMI selecionado para os dados do Experimento 3 após a correção dos autovalores pelos métodos 1, 2 e 3	52
26	Análise de variância conjunta do Experimento 4 com 8 genótipos avaliado em 59 ambientes com 4 blocos	52
27	Decomposição de soma de quadrados de interação do Experimento 4 com 8 genótipos avaliados em 59 ambientes com 4 blocos	52
28	Correção dos autovalores da matriz $(GE)(GE)^t$ do experimento 4 e os autovalores ajustados pela regressão isotônica	53
29	Análise do Teste F para os componentes do Experimento 4 corrigidos pelo método 1	54
30	Análise do Teste F para os componentes do Experimento 4 corrigidos pelo método 2	54
31	Análise do Teste F para os componentes do Experimento 4 corrigidos pelo método 1	55
32	$RMSPD_{PRESS}$ e R_{AMMI} para o melhor modelo AMMI selecionado para os dados do Experimento 4 após a correção dos autovalores pelos métodos 1, 2 e 3	55

1 Introdução

Ao realizar um experimento em um local, com todas as parcelas agrupadas em uma pequena área, os resultados obtidos são válidos somente para a área em questão. Para áreas maiores é necessário realizar experimentos distintos, relativamente simples, onde cada um é válido para uma pequena área, ou bloco (PIMENTEL-GOMES, 2000).

Em muitos casos, o pesquisador está interessado em avaliar o desempenho de vários genótipos (tratamentos) em diversos ambientes (locais e/ou ano). Quando um conjunto de experimentos é planejado para vários locais é necessário considerar o delineamento individual em cada local e a combinação total dos delineamentos com os locais (interação genótipo $\times$ ambiente). Logo, os dados observados podem ser organizados em uma tabela de dupla entrada, colocando, por exemplo, os genótipos nas linhas e os ambientes nas colunas.

Para investigar a existência da interação genótipo $\times$ ambiente, é necessário usar a técnica de análise da variância (ANAVA), mas esta técnica não é suficiente para generalizar os resultados. Embora a interação represente uma das principais dificuldades encontradas pelos melhoristas durante a seleção de material genético, é importante ressaltar que com a exploração desse efeito pode-se chegar a ótimos resultados.

Várias metodologias estatísticas têm sido propostas para a interpretação da interação genótipo $\times$ ambiente e os pesquisadores ainda continuam na busca de uma ferramenta estatística que permita extrair grande parte da informação possível desta fonte de variação. As ferramentas que têm sido utilizadas são os métodos de regressão linear simples (EBERHART; RUSSEL, 1966) e regressão linear múltipla (SILVA; BARRETO, 1985), mas as metodologias baseadas em regressão possuem limitações e têm sido alvo de várias críticas, como no caso em que a linearidade falha.

Na busca de novas ferramentas para o estudo da interação genótipo $\times$ ambiente, o modelo AMMI (GOLLOB, 1968; MANDEL, 1969,1971) vem se destacando e ganhando grande aplicabilidade nos últimos anos (DUARTE; VENCOSKY, 1999). A metodologia AMMI engloba num único modelo a técnica ANAVA, para efeitos principais, e para efeitos multiplicativos (interação), utiliza-se análise de componentes principais (ACP) ou decomposição em valor singular (DVS), aplicado à matriz de interação.

Essas técnicas multivariadas são baseadas no uso dos autovalores e autovetores prove-

nientes da matriz de interação genótipo $\times$ ambiente. Araújo e Dias (2005) verificaram o problema de superestimação e subestimação de autovalores estimados da maneira convencional. Para corrigir esses problemas de estimação de autovalores, Muirhead (1987) apresenta três métodos para corrigir autovalores estimados a partir das matrizes de covariâncias amostral. Este autor alerta que nem sempre essas correções mantêm a ordem decrescente de valores, assim é sugerido que se use regressão isotônica para ordenar esses dados.

Assim, a presente dissertação tem os seguintes objetivos: apresentar três técnicas analíticas para a correção de autovalores com ordenação dos mesmo por regressão isotônica e aplicar estas correções aos modelos AMMI, em estudo de experimentos agronômicos multiambientais; verificar o ganho, em número de repetições ao analisar experimentos utilizando a melhor das três correções para os autovalores; e implementar uma rotina computacional para a análise de dados, segundo modelo AMMI, utilizando as correções para os autovalores sugeridas por Muirhead (1987).

2 Desenvolvimento

2.1 Revisão de Literatura

2.1.1 Interação de genótipos $\times$ ambientes

Em experimentação agrícola, é freqüente a necessidade de análise conjunta de grupos de experimentos. Isto ocorre quando se deseja generalizar resultados para condições gerais de regiões. Nesse caso, uma série de experimentos é conduzida em locais representativos da região e repetidos um número suficiente de vezes de forma que represente as alternativas do clima da região.

No processo de melhoramento genético, a influência que o ambiente exerce sobre cada genótipo tem merecido atenção especial por parte dos melhoristas, visto que a interação entre genótipos e ambientes pode interferir negativamente ou positivamente nos resultados finais (GONZÁLEZ, 1988).

Quando vários cultivares são avaliados em muitos ambientes, pode acontecer de o ambiente influenciar de modo diferente as cultivares em questão, considerando a mesma característica em estudo, ou seja, pode ocorrer uma interação entre os genótipos e os ambientes. Assim, pode-se definir a interação de genótipos com ambientes como sendo o efeito residual de respostas diferentes dos ambientes sobre os genótipos. Hoogerheide (2004) cita que para Chaves (2001)¹, a interação resulta da resposta diferente dos genótipos em ambientes distintos. Essa interação pode ser provocada por fatores fisiológicos, adaptativos e relativos a escalas de medidas da variável, entre outros fatores (CRUZ; CARNEIRO, 2003). Por outro lado Cruz e Regazzi (2001) citam ainda que a interação pode ser atribuída a fatores bioquímicos próprios de cada genótipo cultivado. A interação de genótipos $\times$ ambientes geralmente mostra diferença na expressão final do potencial dos genótipos (EBERHART; RUSSEL, 1966).

Estudos sobre interação genótipos $\times$ ambiente têm proporcionado informações sobre como eliminar as tendências de se superestimarem as variâncias genéticas, sendo que essa superestimação é que leva a discrepância entre as repostas esperadas e obtidas com a seleção, adaptação de cultivares, e avaliação de genótipos e estabilidade de produção (ALLARD, 1971).

Para Allard e Bradshaw (1964) as variáveis ambientais podem ser classificadas em

¹CHAVES, J.L. Interação de cultivares com ambientes. In: NASS, L.L.; VALOIS, A.C.C.; MELO, I.S.; VALADARES, M.C. **Recursos genéticos e melhoramento - plantas**. Rondonópolis: Fundação MT, 2001, p.673-713

dois tipos: previsíveis e imprevisíveis. As variáveis previsíveis seriam as características gerais do clima e solos que ocorrem de maneira sistemática ou que estão sob controle do homem. Já as variáveis imprevisíveis correspondem às flutuações climáticas tais como quantidade e distribuição de chuvas, temperatura e outros fatores que não podem ser controlados pelo homem. Essas variáveis imprevisíveis são as que mais contribuem para a interação (Fehr, 1987).

A interação genótipos $\times$ ambiente é uma fonte de variação fenotípica, que na maioria dos casos é inseparável da variância ambiental (FALCONER, 1987). Na prática, para verificar a significância da interação de genótipos com ambientes, é necessário repetir o experimento várias vezes, pois se o experimento for realizado somente em um ambiente, poderá ocorrer uma superestimação dos ganhos genéticos (CROSSA, 1990).

Com os dados organizados em uma tabela de dupla entrada, a interação entre genótipos $\times$ ambientes é facilmente estudada. O processo usual de investigação da interação entre os genótipos e os ambientes é feito através da análise da variância conjunta (ANAVA), envolvendo vários experimentos, e por essa análise é possível determinar a magnitude da interação, através da razão do Quadrado Médio da interação ($QM_{G \times E}$) pelo Quadrado Médio do Erro Médio (QM_{EM}). Estatisticamente, a interação genótipos $\times$ ambientes são detectadas como respostas diferentes e significantes dos genótipos entre os ambientes. A detecção de significância para a interação não esclarece as conclusões que se possa ter sobre o melhoramento, sendo necessário um estudo detalhado deste componente de variação.

O estudo da interação genótipo $\times$ ambiente é de interesse especial desde que o estudo possa ser útil na identificação da subdivisão de certa região dentro de diferentes sub-regiões (regionalização), objetivando indicar áreas onde essas possam expressar o máximo que as condições ambientais particulares permitam, com respeito a respostas de genótipos. Possibilita ainda explorar efeitos específicos de adaptação para determinadas regiões; alternativamente, isto pode promover formas de procedimentos para a escolha de um local do ponto de vista da adaptação e também características adaptativas podem ser simultaneamente identificadas (GONZÁLEZ, 1988; ANNICCHIARICO, 1997b).

Somente a verificação da interação não é suficiente, mas que se deve considerar a sua natureza (VENCOVSKY; BARRIGA, 1992). Assim, a natureza pode ser simples e complexa. A interação de natureza simples indica a presença de genótipos adaptados em um grande número de ambientes, sendo possível fazer uma recomendação generalizada de cultivares. A interação com

natureza complexa mostra que existem genótipos adaptados a apenas alguns ambientes, o que traz uma complicação ao pesquisador, quando da recomendação de cultivar.

A existência de interação genótipo $\times$ ambiente, produz uma barreira de dificuldades aos melhoristas na identificação de genótipos superiores, tanto no processo de seleção quanto no processo de recomendação de cultivares. Essa interação indica que o comportamento dos genótipos nos experimentos depende principalmente das condições ambientais a que são submetidos. Assim a resposta obtida de um genótipo, em comparação a outro, é variável, sendo que essas variações se apresentam devido a mudança de ambientes (OLIVEIRA; DUARTE; PINHEIRO, 2003; KANG, 1998).

Na fase inicial de seleção dos programas de melhoramento, as avaliações dos genótipos, são realizadas em um só local e conseqüentemente, as estimativas da variância genética são acrescidas do componente de variância da interação ($G \times E$). Diante disso, os ganhos genéticos previstos com a seleção não são reais. Na fase seguinte, os ensaios são conduzidos em vários ambientes que podem ser a combinação de vários locais, anos e/ou épocas. Desta forma é possível isolar o componente de variância da interação ($G \times E$), porém a intensidade de seleção é baixa fazendo com que o ganho genético seja de pequena magnitude. Assim, os efeitos principais de genótipos e ambientes não podem ser interpretados de forma isolada e deve-se considerar interação ($G \times E$) (DUARTE; VENCOSKY, 1999; KANG; MAGARI, 1996; MANDEL, 1971).

Várias metodologias têm sido propostas no sentido de entender melhor o efeito da interação genótipo $\times$ ambiente. Algumas dessas propostas são: zoneamento ecológico ou estratificação de ambientes, ou seja, identificar regiões ou sub-regiões onde o efeito da interação seja não significativa, pode levar a identificação de genótipos que se adaptam a ambientes específicos e ainda identificar genótipos com uma ampla adaptação ou estabilidade (RAMALHO; SANTOS; ZIMMERMANN, 1993).

Assim, a interação de genótipos $\times$ ambientes deve ser considerada, não como um problema ou um fator indesejável, cujo efeito deve ser minimizado, mas deve ser enfrentada como um fenômeno biológico natural, que deve ser bem conhecido para melhor aproveitá-lo no processo de seleção (CHAVES, 2001). Logo os genótipos que interagem positivamente com os ambientes podem fazer a diferença entre um bom e um ótimo cultivar (DUARTE; VENCOSKY, 1999).

2.1.2 Modelos AMMI

São muitas as metodologias de análise e interpretação para a interação genótipo $\times$ ambiente proveniente de um grupo de cultivares testados em vários ambientes. Para Vencovsky e Barriga (1992) a diferença entre eles origina-se nos próprios conceitos de estabilidade e nos procedimentos biométricos de medir a interação cultivares e ambientes.

A metodologia AMMI (“additive main effects and multiplicative interaction model”), como o próprio nome diz é um método uni-multivariado, que engloba uma análise de variância para os efeitos principais, que são os efeitos dos genótipos (G) e os ambientes (E) e ainda leva em conta uma análise multivariada dos efeitos provocados pela interação genótipo $\times$ ambiente. Uma breve análise histórica é apresentada por Gauch Júnior (1992), considerando os métodos que fazem parte da análise AMMI: A análise de componentes principais (ACP) e a análise de variância (ANAVA).

Foi introduzido por Gollob (1968) e Mandel (1969,1971) um modelo (FANOVA) com fator analítico para estudar a interação em tabelas de dupla entradas completas. Zobel; Wright e Gauch Júnior (1988), Gauch (1988); Gauch Júnior e Zobel (1988) renomearam o modelo FANOVA como “Additive Main Effects and Multiplicative Interactions (AMMI) model”.

Ambas as análises (ANAVA e ACP) buscam resumir os dados contidos em uma matriz de interação genótipo $\times$ ambiente em um modelo. A única diferença do ponto de vista conceitual é que na ANAVA, as estimativas dos efeitos dos genótipos e dos ambientes são somadas à média geral, enquanto que na ACP os efeitos são multiplicados entre si e depois são somados à média geral (GAUCH JÚNIOR, 1992).

Para Miranda (2004) ACP foi criada por Pearson em 1902, enquanto que a ANAVA foi proposta por Fisher em 1918 e apesar da ACP ser mais antiga e ser considerada mais eficiente, é a ANAVA que se tornou mais popular devido à simplicidade dos cálculos e à fácil interpretação do modelo aditivo. Já a ACP foi pouco utilizada devido à exigência de grandes recursos computacionais, ausente quando a teoria foi proposta e pela dificuldade de interpretação do modelo multiplicativo.

Em relação ao modelo AMMI, Gauch Júnior (1992) faz algumas considerações importantes, dizendo considerar que o modelo AMMI foi proposto simultaneamente por Pike e Silverberg (1952) e Williams (1952), sob a forma de soma de funções mais o produto de funções, onde essas

funções podem ser substituídas por números. Estes autores utilizaram a ACP apenas na matriz de resíduos de interação genótipos $\times$ ambiente resultante da ANAVA.

É apresentado por Gabriel (1971) o gráfico “biplot” para visualização de qualquer matriz por uma matriz de aproximação ou que possa ser resumida em uma matriz de posto dois, com especial importância na visualização de resultados originados da ACP.

O primeiro a utilizar a metodologia de gráficos “biplot” em dados de produtividade, enfatizando a possibilidade de utilizar a técnica na interpretação dos dados foi Kempton (1984). Este autor apresenta a utilização e interpretação de “biplots” para os modelos de ANAVA, RL (regressão linear), ACP, e AMMI e considera o “biplot” uma interessante ferramenta para a apresentação dos resultados das diversas técnicas, com especial interesse para a análise da interação genótipo $\times$ ambiente.

Em Gauch Júnior e Zobel (1996) é feita uma revisão da metodologia AMMI no período de 1992 a 1996, na qual mostram ser útil para apresentar o melhor entendimento de interações complexas genótipo $\times$ ambiente e comentam também sobre a análise gráfica “biplot”, a capacidade do modelo AMMI em construir um modelo rico em padrão, deixando o resíduo rico em ruído, o que faz com que o modelo tenha uma maior precisão. Para estes autores a utilização do modelo AMMI significa um ganho de precisão equivalente ao aumento de 5 a 20 repetições.

Comparando a ANAVA, a ACP, a RL (metodologias tradicionais) e o modelo AMMI, Zobel; Wright e Gauch Júnior (1988) verificaram que a ANAVA falhou em detectar a significância da interação genótipo $\times$ ambientes, já a ACP falhou em identificar e separar os componentes principais de genótipos e de ambientes, a RL (a metodologia mais utilizada) falhou em reter apenas uma pequena parcela da soma de quadrados da interação genótipo $\times$ ambiente, misturando o efeito da interação genótipo $\times$ ambiente com os efeitos principais e a análise AMMI mostrou uma alta e significativa interação de genótipo $\times$ ambiente com sentido agrônomo, permitindo separar os ambientes quanto a latitude e os genótipos quanto à precocidade.

Analizando quatro grupos de dados de diferentes culturas, Annicchiarico (1997a) compara dois procedimentos baseados em validação cruzada para testar os eixos dos componentes principais da interação genótipo $\times$ ambiente. O procedimento de validação cruzada consiste em dividir o conjunto de dados em duas partes de forma aleatória, sendo que uma parte é destinada para ajustar o modelo e a segunda parte para validá-lo. Assim, o modelo AMMI é obtido com

dados destinados para ajustar o modelo e a soma de quadrados residual (RMSPD) é determinada nos dados destinados para validação. Para experimentos realizados em vários ambientes, os ajustes devem ser feitos para as repetições dentro dos ambientes, sendo que o melhor modelo será aquele que apresentar menor RMSPD (DIAS; KRZANOWSKI, 2003).

É feita uma comparação, Annicchiarico (1997a), do teste F de Gollob (1968), F_{GH2} e F_R ambos desenvolvidos por Cornelius; Seyedsadr e Crossa (1992) que são utilizados para selecionar uma série de termos multiplicativos estatisticamente significantes, ou seja, o teste F trunca o modelo completo em certo ponto. Os resultados não apresentaram concordância na seleção de um modelo truncado para um particular conjunto de dados e o método de Gollob é considerado adequado apenas para pequenos conjuntos de dados.

Utilizando estimadores de encolhimento (“shrinkage”) para termos multiplicativos do modelo AMMI, Cornelius e Crossa (1999) verificaram que estes estimadores são sempre melhores que aqueles baseados em modelos truncados ajustados por mínimos quadrados e validação cruzada. Os autores ainda verificaram que estes estimadores são ao menos tão eficientes quantos os BLUPs (“Best Linear Unbiased Prediction”). Diante dessa superioridade sobre os outros estimadores, as estimativas encolhidas produziram predições muito mais acuradas e ainda se torna desnecessária a validação cruzada e os testes de hipóteses. Já os melhores graus de liberdade são aqueles obtidos pelo método de Gollob (1968), facilmente obtidos e estão bem próximos dos graus de liberdade obtidos por simulação.

Outro método de “shrinkage” para os estimadores dos termos multiplicativos do modelo AMMI é proposto por Moreno-González; Crossa; Cornelius (2003a) e em Moreno-González; Crossa; Cornelius (2003b) é feita uma comparação deste com aquele proposto por Cornelius e Crossa (1999). Os autores concluíram que ambos os métodos são similares e que a união dos dois métodos (MORENO-GONZÁLEZ; CROSSA; CORNELIUS, 2003b e CORNELIUS; CROSSA, 1999) finalmente estabelece que para separar a resposta padrão do ruído, em estimativas de respostas médias, os modelos que utilizam estimadores de encolhimento para os fatores são superiores aos modelos AMMI ajustados por mínimos quadrados parcimoniosos.

Assim, a abordagem AMMI não busca recuperar toda $SQ_{G \times E}$, mas apenas uma parcela que melhor representa, determinada por genótipos e ambientes. Essa parcela é formada por eixos que captam, sucessivamente, porções cada vez menores da variação presente na matriz de interação GE . Por isso, o método AMMI é considerado um modelo capaz de separar a resposta

padrão e ruído da $SQ_{G \times E}$ (WEBER; WRICKE; WESTERMANN, 1996).

2.1.3 Autovalores e Autovetores de uma matriz

Seja $\mathbf{D} = \mathbf{X}\mathbf{X}^t$, uma matriz de ordem p . Esta matriz poder ser vista como um operador linear (transformação linear) que leva um vetor $\mathbf{u}$ não nulo, em um vetor $\mathbf{z}$ pertencente ao mesmo espaço de $\mathbf{u}$, como segue:

$$\mathbf{D}\mathbf{u} = \mathbf{z}. \quad (1)$$

Seja $\mathbf{z}$ um vetor proporcional a $\mathbf{u}$, com fator de proporcionalidade λ , então defini-se λ como sendo o autovalor ou o valor principal de $\mathbf{D}$ associado ao autovetor $\mathbf{u}$ (BOLDRINI et al., 1980). Dessa forma a expressão (1) torna-se:

$$\mathbf{D}\mathbf{u} = \lambda\mathbf{u}. \quad (2)$$

O sistema (2) é conhecido como o “problema do autovalor” ou autoestrutura de $\mathbf{D}$ (LANCZOS, 1961) e possui duas incógnitas a determinar: λ e $\mathbf{u}$. A solução é obtida com a solução do sistema homogêneo dado por:

$$(\mathbf{D} - \lambda\mathbf{I})\mathbf{u} = \mathbf{0}. \quad (3)$$

sendo que $\mathbf{0}$ é um vetor nulo e $\mathbf{I}$ é uma matriz identidade de ordem p . O sistema (3) admite solução não trivial se e somente se

$$s(\lambda) = \det(\mathbf{D} - \lambda\mathbf{I}) = 0 \quad (4)$$

A função $s(\lambda)$ é um polinômio de grau p conhecido como polinômio característico da transformação e suas raízes são os autovalores de $\mathbf{D}$.

Após a determinação dos autovalores λ_i , com $i = 1, 2, \dots, p$, retorna-se à expressão (2) e calcula-se o autovetor associado a cada autovalor λ_i .

As principais propriedades dos autovalores e autovetores de matrizes simétricas e de covariância são:

- i) Se $\mathbf{D}$ é uma matriz simétrica, seus autovalores são reais;
- ii) Se $\mathbf{D}$ é uma matriz simétrica e seus autovalores são distintos, então os autovetores associados são ortogonais entre si e formam uma base do espaço vetorial;
- iii) Se $\mathbf{D}$ é uma matriz de covariâncias e, portanto, positiva definida, então seus autovalores são sempre positivos;
- iv) Se λ_i , com $i = 1, 2, \dots, p$ são os autvalores de $\mathbf{D}$, então o $Tr(\mathbf{D}) = \sum_{i=1}^p \lambda_i$ e $det(\mathbf{D}) = \prod_{i=1}^p \lambda_i$;
- v) Seja $\mathbf{X}$ uma matriz retangular de ordem $n \times m$, cujo posto seja $p < \min(n, m)$. Supondo $n < m$, então, os p autovalores não nulos de $\mathbf{X}\mathbf{X}^t$ são iguais aos p autovalores não nulos de $\mathbf{X}^t\mathbf{X}$. Os demais autovalores são iguais a zero.

2.1.3.1 Estimação de autovalores e distribuição amostral

Seja $\mathbf{\Sigma}$ uma matriz simétrica e positiva definida de ordem p . Pode-se encontrar um conjunto de autovalores $\lambda_1 > \lambda_2 > \dots > \lambda_p > 0$. De forma análoga, é possível encontrar os autovalores de uma matriz $\mathbf{S}$ que estima $\mathbf{\Sigma}$ e calcular os autovalores $\hat{\lambda}_1 > \hat{\lambda}_2 > \dots > \hat{\lambda}_p > 0$ de $\mathbf{S}$. Como $\mathbf{S}$ é uma estimativa de $\mathbf{\Sigma}$, tem-se que cada $\hat{\lambda}_i$, com $i = 1, 2, \dots, p$, é uma estimativa de λ_i (Muirhead, 1987).

Tomando uma amostra aleatória $X_1, X_2, \dots, X_m$, proveniente de uma população normal multivariada de dimensão m com matriz de covariâncias populacional $\mathbf{\Sigma}$ (que possui autovalores $\lambda_1 > \lambda_2 > \dots > \lambda_p > 0$) e o vetor de médias $\boldsymbol{\mu}$ conhecido é possível obter a matrix $\mathbf{S}$ (que possui autovalores $\hat{\lambda}_1 > \hat{\lambda}_2 > \dots > \hat{\lambda}_p > 0$). Johnson e Wichern (1998) e Anderson (1958) assumem para uma amostra grande, que cada $\hat{\lambda}_i$ têm uma distribuição independente. Estes autores declaram também que cada $\hat{\lambda}_i$ têm uma distribuição assintótica normal, ou seja, $\hat{\lambda}_i \stackrel{a}{\sim} N(\lambda_i, \frac{2\lambda_i^2}{n})$. Araújo e Dias (2005) confirmam essa declaração, simulando amostras multivariadas e depois verificando a aderência da distribuição dos autovalores à distribuição normal especificada.

2.1.3.2 Viés e Correção de autovalores

Seja $\mathbf{S}$, com autovalores $\hat{\lambda}_1 > \hat{\lambda}_2 > \dots > \hat{\lambda}_p$, a estimativa de uma matriz de covariância populacional $\mathbf{\Sigma}$ de ordem p , com autovalores $\lambda_1 > \lambda_2 > \dots > \lambda_p$. Muirhead (1987) afirma que esses autovalores são viesados. Araújo e Dias (2005), utilizando técnicas de simulação multivariadas, confirmaram isso e concluíram que os maiores $\hat{\lambda}_i$ superestimam os respectivos λ_i , enquanto que os menores $\hat{\lambda}_i$ subestimam os respectivos λ_i .

Seja $\mathbf{X}$ uma matriz de ordem $n \times m$, onde cada linha representa um vetor normal multivariado de dimensão m com matriz de covariâncias $\mathbf{\Sigma}$ e vetor de médias $\boldsymbol{\mu}$. Segundo Johnstone (2001), se $\boldsymbol{\mu} = 0$, a matriz $\mathbf{S} = \mathbf{X}^t \mathbf{X}$ é uma estimativa de $\mathbf{\Sigma}$ e tem distribuição de Wishart com n graus de liberdade e matriz de covariâncias $\mathbf{\Sigma}$, isto é: $\mathbf{S} \sim W_m(n, \mathbf{\Sigma})$

Para uma matriz $\mathbf{S} \sim W_m(n, \mathbf{\Sigma})$, Muirhead (1987) apresenta três métodos de correção (26), (27) e (28) (página 33) com a proposta de corrigir esse viés dos autovalores.

Ao corrigir os autovalores pelos métodos (26) e (27) tem-se que a restrição ($\phi_1 \geq \phi_2 \geq \dots \geq \phi_p$) nem sempre é verificada, onde ϕ_i é a correção do autovalor $\hat{\lambda}_i$, com $i = 1, 2, \dots, p$. Para ordenar essas correções é necessário utilizar a regressão isotônica, que passará a ser discutida.

2.1.4 Regressão Isotônica

A técnica de regressão ajusta curvas ou funções para um conjunto de pontos e existem vários métodos estatísticos que verificam a qualidade do ajuste.

Em várias situações de pesquisa acredita-se, devido a informações a “priori”, que os dados devem apresentar uma característica relativa a uma ordenação. Assim, procura-se um modelo que preserve essa característica de ordem. O modelo selecionado será escolhido dentre vários modelos que considere essa ordem dos dados e esse método é denominado *regressão isotônica*. O termo isotônica (ou monotônica) refere-se ao fato de a variável resposta aumentar de acordo com o aumento da variável independente. Se a variável resposta apresentar uma característica decrescente com o aumento da variável preditora, pode-se usar o termo antitônica. Mas esses termos isotônica, monotônica e antitônica possuem o mesmo sentido.

Dentre os diversos tipos de conjuntos existentes, a regressão isotônica é feita em dois

conjuntos particulares. Esses conjuntos são: os conjuntos com uma simples ordenação de dados e os conjuntos com uma quase ordenação de dados.

2.1.4.1 Regressão Isotônica: Um caso simples de dados ordenados

Seja $X = \{x_1 \lesssim x_2 \lesssim \dots \lesssim x_p\}$ um conjunto finito. A relação binária ($\lesssim$), entre os elementos de X , é uma relação de ordem simples se são satisfeitas as seguintes propriedades:

1. é reflexiva: $x \lesssim x$ para todo $x \in X$;
2. é transitiva: $x, y, z \in X$, $x \lesssim y$ e $y \lesssim z$ então $x \lesssim z$;
3. é anti-simétrica: $x, y \in X$, $x \lesssim y$ e $y \lesssim x$ então $x = y$; e
4. todo e qualquer elemento de X é comparável: $x, y \in X$ implica que $x \lesssim y$ ou $y \lesssim x$.

Assim, Robertson; Wright e Dykstra (1998) apresentam a seguinte definição para regressão isotônica, no caso de uma ordenação simples dos dados.

Seja $X = \{x_1, x_2, \dots, x_p\}$ um conjunto finito com uma ordem simples $x_1 \lesssim x_2 \lesssim \dots \lesssim x_p$. Uma função f , em X , é isotônica se $f(x_1) \leq f(x_2) \leq \dots \leq f(x_p)$, sendo que ($\lesssim$) indica a relação existente no conjunto da variável independente e ($\leq$), a relação existente no conjunto da variável dependente. Seja também w uma função peso, positiva e definida em X .

Suponha que t seja uma função dada em X . Uma função t^* é uma regressão isotônica de t com peso w se e somente se t^* é uma função isotônica e minimiza

$$\sum_{x \in X} [t(x) - f(x)]^2 w(x) \quad (5)$$

dentro de todas classes de funções isotônicas f em X .

Assim, para uma função arbitrária t , definida em X , seja t^* a projeção de mínimos quadrados de t dentro da coleção de funções isotônicas. Especificamente, t^* é uma solução de mínimos quadrados restritos para a expressão (5).

Suponha amostras aleatórias de p populações normais com médias $\mu(x_1), \mu(x_2), \dots, \mu(x_p)$ e variância comum σ^2 . Para a i -ésima amostra, seja n_i a notação

para o seu tamanho amostral, sejam $Y_{i1}, Y_{i2}, \dots, Y_{in_i}$ as observações e $\bar{Y}_i = \sum_{j=1}^{n_i} \frac{Y_{ij}}{n_i}$ o estimador de máxima verossimilhança (EMV), sem restrição de $\boldsymbol{\mu} = (\mu(x_1), \mu(x_2), \dots, \mu(x_p))$, o vetor de médias populacionais. Suponha também, que é conhecido que $\mu(x_1) \leq \mu(x_2) \leq \dots \leq \mu(x_p)$. Pode-se querer estimar essas médias e também satisfazer essa restrição, mas pela variabilidade amostral, a amostra de médias $\bar{Y}_1, \bar{Y}_2, \dots, \bar{Y}_p$, pode não estar ordenada, ou seja, não está obedecendo à restrição das médias populacionais.

Para obter as estimativas que satisfazem esta restrição, toma-se o negativo do logaritmo da função de verossimilhança. Percebe-se que o EMV restrito minimiza $\sum_{i=1}^p \sum_{j=1}^{n_i} [Y_{ij} - \mu(x_i)]^2$ sujeito a $\mu(x_1) \leq \mu(x_2) \leq \dots \leq \mu(x_p)$. Somando e subtraindo $\bar{Y}_i$, podemos reescrever esse somatório da seguinte maneira:

$$\sum_{i=1}^p \sum_{j=1}^{n_i} (Y_{ij} - \bar{Y}_i)^2 + 2 \sum_{i=1}^p \sum_{j=1}^{n_i} (Y_{ij} - \bar{Y}_i) [\bar{Y}_i - \mu(x_i)] + \sum_{i=1}^p [\bar{Y}_i - \mu(x_i)]^2 n_i.$$

Entretanto, a primeira soma não envolve $\mu(x_i)$ e a segunda soma é zero. Assim o EMV restrito minimiza $\sum_{i=1}^p [\bar{Y}_i - \mu(x_i)]^2 n_i$ sujeito à condição de $\mu(x_1) \leq \mu(x_2) \leq \dots \leq \mu(x_p)$. Portanto, em um conjunto $X = \{x_1, x_2, \dots, x_p\}$ com $x_1 \lesssim x_2 \lesssim \dots \lesssim x_p$, t e w são definidas em X por $t(x_i) = \bar{Y}_i$ e $w(x_i) = n_i$, então está claro que o EMV restrito de $\boldsymbol{\mu} = (\mu(x_1), \mu(x_2), \dots, \mu(x_p))$ é fornecido pela regressão isotônica t^* de t com peso w .

2.1.4.2 O algoritmo PAVA

Um algoritmo amplamente utilizado para calcular a regressão isotônica para uma ordem simples é o PAVA (*pool-adjacent-violators algorithm*). Robertson; Wright e Dykstra (1988) citam que Ayer et al. (1955)² foram os primeiros a publicarem o algoritmo e apresentam o processo de isotonização que se segue.

O algoritmo começa com a função $t(x)$. Se $t(x)$ é isotônica, então $t(x) = t^*(x)$. Caso contrário, deve existir algum índice i tal que $t(x_{i-1}) > t(x_i)$. Estes dois valores são trocados pelas suas médias ponderadas ($Av(i-1, i)$), isto é:

² AYER, M.; BRUNK, H.; EWING, G.; REID, W.; SILVERMAN, E. An empirical distribution function for sampling with incomplete information. **Annals of Mathematical Statistics**, Ann Arbor, v.26, n.4, p.641-647, 1955.

$$Av(i-1, i) = \frac{t(x_{i-1})w(x_{i-1}) + t(x_i)w(x_i)}{w(x_{i-1}) + w(x_i)}. \quad (6)$$

Os pesos $w(x_{i-1})$ e $w(x_i)$, são trocados pela soma desses, isto é, $[w(x_{i-1}) + w(x_i)]$.

Se esse novo conjunto, com $p-1$, valores, é isotônico ($t(x_1) \leq \dots \leq t(x_{i-2}) \leq Av(i-1, i) \leq \dots \leq t(x_p)$) então $t^*(x_{i-1}) = t^*(x_i) = Av(i-1, i)$ e $t^*(x_j) = t(x_j)$, com $j = 1, 2, \dots, (p-1)$.

Caso esse novo conjunto, de valores não seja isotônico, esse processo é repetido, usando os novos valores e novos pesos, até obter um conjunto de valores isotônico, sendo que o valor $t^*(x_i)$ é a média ponderada do bloco em que x_i está contido.

2.1.4.3 Regressão isotônica em um conjunto quase-ordenado

A relação binária $\lesssim$ em um conjunto $X = \{x_1 \lesssim x_2 \lesssim \dots \lesssim x_p\}$ é uma ordem parcial se são satisfeitas as propriedades de reflexão, transitividade e anti-simetria. Um conjunto X , com a relação binária $\lesssim$, ainda pode ser um conjunto quase-ordenado se os elementos deste conjunto satisfizerem somente as propriedades de reflexão e transitividade, não sendo necessário que elementos sejam anti-simétricos e nem que sejam comparáveis. Assim, todo conjunto com ordem simples é um conjunto com ordem parcial e todo conjunto com ordem parcial é um conjunto com quase-ordem.

Uma função f de valor real em X é isotônica com elementos quase-ordenados ($\lesssim$) em X se $x, y \in X$ então $f(x) \leq f(y)$. Assim, seja t uma função dada em X e w uma outra função positiva em X . Uma função isotônica t^* é uma regressão isotônica de t com peso w se e somente se

$$\sum_{x \in X} [t(x) - t^*(x)]^2 w(x) \leq \sum_{x \in X} [t(x) - f(x)]^2 w(x) \quad (7)$$

para todas funções f em X que são isotônicas.

2.1.4.4 Regressão isotônica generalizada

Suponha que Φ seja uma função convexa a qual é finita em um intervalo I contendo uma gama de funções t em X e tem o valor $+\infty$ fora do intervalo I . Então, é bem conhecido

que Φ é contínua em I , em I^0 (conjunto dos extremos do intervalo I) e, de fato, Φ é diferenciável pela direita e pela esquerda de cada ponto em I^0 . Seja ϕ um função não decrescente em I que coincida com alguma determinação de uma derivada de Φ em I^0 , ou seja, para cada $x \in I$, $\phi(x)$ é um número entre a derivada à esquerda de Φ em x e à derivada a direita de Φ em x . O valor de ϕ à esquerda do maior ponto de I pode ser $-\infty$ e este valor pode ser $+\infty$ à direita do menor ponto de I .

Para cada $u, v \in I$, define-se uma função $\Delta_\Phi(u, v)$ por:

$$\Delta_\Phi(u, v) = \Phi(u) - \Phi(v) - (u - v)\phi(v)$$

com $\Delta_\Phi(u, u) = 0$ e $\Delta_\Phi(u, v) = +\infty$, se u ou v não pertencer a I .

Se f é uma função isotônica em X e se as possíveis f estiverem dentro de I então

$$\sum_{x \in X} \Delta_\Phi[t(x), f(x)]w(x) \geq \sum_{x \in X} \Delta_\Phi[t(x), t^*(x)]w(x) + \sum_{x \in X} \Delta_\Phi[t^*(x), f(x)]w(x). \quad (8)$$

Conseqüentemente t^* minimiza

$$\sum_{x \in X} \Delta_\Phi[t(x), f(x)]w(x) \quad (9)$$

na classe de todas as funções isotônicas f existentes dentro de I e maximiza

$$\sum_{x \in X} \{\Phi[f(x)] + [t(x) - f(x)]\phi[f(x)]\}w(x). \quad (10)$$

2.1.5 Decomposição em valores singulares

A decomposição em valores singulares (“singular value decomposition”, DVS) é um dos resultados mais importantes da álgebra linear, tanto computacional quanto teórico, que é aplicada em diversas áreas de conhecimento. A DVS é a base dos métodos mais precisos para a resolução de problemas de mínimos quadrados, para a determinação de posto de matrizes, do espaço-imagem ($R(\cdot)$) e do espaço nulo ($N(\cdot)$) de matrizes e de solução de vários problemas envolvendo normas euclidianas($||\cdot||$).

Mandel (1982) apresenta um resumo da teoria sobre DVS para aplicação em análise de regressão. Para qualquer matriz $\mathbf{X}$ de ordem $n \times m$, com posto $p < \min(n, m)$, este autor diz que é possível expressar cada elemento de $\mathbf{X}$ da seguinte maneira:

$$\begin{aligned}
x_{ij} &= \lambda_1 u_{1i} v_{1j} + \lambda_2 u_{2i} v_{2j} + \dots + \lambda_p u_{pi} v_{pj} \\
&= \sum_{k=1}^p \lambda_k u_{ki} v_{kj},
\end{aligned} \tag{11}$$

em que $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_p$. Esta é a definição de DVS da matriz $\mathbf{X}$.

Os p vetores $\mathbf{u}_i$ são ortogonais a cada vetor $\mathbf{u}_j$, para $i \neq j$ e $i, j = 1, 2, \dots, p$, assim como os p vetores $\mathbf{v}_i$. Além disso, considera-se que cada um desses vetores tem tamanho unitário (vetores normalizados), ou seja

$$\sum_i \mathbf{u}_{ki} = \sum_j \mathbf{v}_{kj} = 1, \quad \forall k. \tag{12}$$

Em notação matricial, escreve-se:

$${}_n \mathbf{X}_m = {}_n \mathbf{U}_p \mathbf{\Lambda}_p \mathbf{V}_m. \tag{13}$$

A matrix $\mathbf{\Lambda}$ é diagonal e todos os λ_k são positivos. As colunas da matriz $\mathbf{U}$ são os vetores $\mathbf{u}$ e as colunas de $\mathbf{V}$ são os vetores $\mathbf{v}$ da equação (11) e (13). A ortogonalidade e a normalidade dos vetores $\mathbf{u}$ e $\mathbf{v}$ implicam nas seguintes condições, na forma matricial:

$$\mathbf{U}^t \mathbf{U} = \mathbf{I} \tag{14}$$

$$\mathbf{V} \mathbf{V}^t = \mathbf{I} \tag{15}$$

sendo que $\mathbf{I}$ indica uma matriz identidade de ordem p . Os autovalores λ_k são obtidos, extraindo a raiz quadrada dos autovalores diferentes de zero da matriz $\mathbf{X}^t \mathbf{X}$ ou a raiz quadrada dos autovalores diferentes de zero da matriz $\mathbf{X} \mathbf{X}^t$. As colunas de $\mathbf{U}$ são os autovetores da matriz $\mathbf{X} \mathbf{X}^t$, ordenadas de acordo com a magnitude dos autovalores e as linhas de $\mathbf{V}^t$ são os autovetores da matriz $\mathbf{X}^t \mathbf{X}$, ordenadas de acordo com a magnitude dos autovalores.

2.2 Material e Métodos

2.2.1 Características dos dados

Os dados a serem utilizados são os mesmos utilizados por Cornelius e Crossa (1999) e Dias e Krzanowski (2003). Foram obtidos pelo CIMMYT (Centro Internacional de Mejoramiento de Maiz y Trigo) em experimentos realizados em vários países, caracterizando assim, experimentos multi-ambientais. Foram utilizados genótipos de milho e trigo sendo que em todos os experimentos utilizou-se o delineamento aleatorizado em blocos. Cada experimento tem a seguinte descrição:

Experimento 1: 20 genótipos de trigo, sendo que um genótipo é do tipo trigo “durum” e os outros 19 são do tipo trigo “bread”. Cada genótipo foi avaliado em 34 ambientes com 4 blocos;

Experimento 2: 16 genótipos de milho em 26 ambientes, com 4 blocos;

Experimento 3: Em cada um dos 20 ambientes, foram conduzidos experimentos com 9 genótipos de milho e com 4 blocos;

Experimento 4: 59 ambientes, 8 genótipos de milho e em cada ambiente teve 4 blocos;

2.2.2 Análises de variância

2.2.2.1 Análise de variância em cada ambiente

Inicialmente, no j -ésimo ambiente com $j = 1, 2, \dots, e$, tem-se um experimento aleatorizado em blocos, então faz-se uma análise de variância (ANAVA), considerando o efeito dos genótipos como fixo, de acordo com o seguinte modelo matemático:

$$Y_{ih} = \mu + g_i + b_h + \varepsilon_{ih} \quad (16)$$

sendo que:

Y_{ih} : é a resposta do i -ésimo genótipo no h -ésimo bloco, com $i = 1, 2, \dots, g$ e $h = 1, 2, \dots, r$;

μ : é uma constante, geralmente a média geral;

g_i : é o efeito do i -ésimo genótipo;

b_h : é o efeito do h -ésimo bloco;

ε_{ih} : é o erro experimental associado ao i -ésimo genótipo no h -ésimo bloco, assumido ser independente e $\varepsilon \sim N(0, \sigma^2)$.

Na Tabela 1 tem-se o esquema da análise de variância para o modelo (16).

Tabela 1 - Esquema da análise de variância para experimento de um grupo de genótipos avaliados em um local em r blocos

Fonte de Variação	Graus de liberdade (GL)	Quadrado Médio (QM)
Blocos (B)	$(r - 1)$	QM_B
Genótipos (G)	$(g - 1)$	QM_G
Resíduos (Res)	$(g - 1)(r - 1)$	QM_{Res}
Total	$(gr - 1)$	

No caso de existirem r repetições em experimentos inteiramente aleatorizado, tem-se o seguinte modelo matemático:

$$Y_{ih} = \mu + g_i + \varepsilon_{ih} \quad (17)$$

sendo que:

Y_{ih} : é a resposta do i -ésimo genótipo na h -ésima repetição, com $i = 1, 2, \dots, g$ e $h = 1, 2, \dots, r$;

μ : é uma constante, geralmente a média geral;

g_i : é o efeito do i -ésimo genótipo;

ε_{ih} : é o erro experimental associado ao i -ésimo genótipo na h -ésima repetição, assumido ser independente e $\varepsilon \sim N(0, \sigma^2)$.

Tem-se na Tabela 2 o esquema da análise de variância para o modelo (17).

Tabela 2 - Esquema da análise de variância para experimento de um grupo de genótipos avaliados em um local em r repetições

Fonte de Variação	Graus de liberdade (GL)	Quadrado Médio (QM)
Genótipos (G)	$(g - 1)$	QM_G
Resíduos (Res)	$g(r - 1)$	QM_{Res}
Total	$(gr - 1)$	

2.2.2.2 Análise de variância conjunta

Com o objetivo de verificar se existe a interação entre genótipos e ambiente, realiza-se uma análise de variância conjunta que envolve o estudo de todos os génotipos em todos os ambientes, sendo que em cada ambiente teve-se um delineamento aleatorizados em blocos. Hill e Rosenberger (1985) e Stroup e Mulitze (1991) mostraram que pode ser preferível, em termos de predição, assumir que o efeito dos genótipos como aleatório e, fixo, o efeito dos ambientes, obtendo o efeito da interação genótipo $\times$ ambiente como aleatório. Os dados serão representados pelo seguinte modelo matemático:

$$Y_{ijh} = \mu + g_i + e_j + (ge)_{ij} + b_{h(j)} + \varepsilon_{ijh} \quad (18)$$

sendo que:

Y_{ijh} : é o valor observado do i -ésimo genótipo no j -ésimo ambiente e no h -ésimo bloco, com $i = 1, 2, \dots, g$, $j = 1, 2, \dots, e$ e $h = 1, 2, \dots, r$;

μ : é uma constante, geralmente a média;

g_i : é o efeito do i -ésimo genótipo;

e_j : é o efeito do j -ésimo ambiente;

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

$b_{h(j)}$: é o efeito do h -ésimo bloco dentro j -ésimo ambiente;

ε_{ijh} : é erro experimental associado ao i -ésimo genótipo, no j -ésimo ambiente e no h -ésimo bloco assumido ser independente e $\varepsilon_{ijh} \sim N(0, \sigma^2)$.

Tabela 3 - Esquema da análise de variância para experimentos de um mesmo grupo de genótipos avaliado em e locais com r blocos

Fonte de Variação	Graus de liberdade (GL)	Quadrado Médio (QM)
Blocos d. ambientes (B d. E)	$e(r - 1)$	$QM_{Bd.A}$
Genótipos (G)	$(g - 1)$	QM_G
Ambientes (E)	$(e - 1)$	QM_E
Interação ($G \times E$)	$(g - 1)(e - 1)$	$QM_{G \times E}$
Resíduo(<i>Res</i>)	$e(g - 1)(r - 1)$	QM_{EM}
Total	$(ger - 1)$	

Na Tabela 3 apresenta-se o esquema da análise de variância para o modelo (18).

Um caso particular e mais simples do modelo (18) é quando o delineamento dentro de cada ambiente for inteiramente ao acaso com r repetições. Assim, é possível encontrar as médias das r repetições, para cada combinação de genótipos e ambientes, neste caso, a ANAVA conjunta é estudada sobre o seguinte modelo:

$$Y_{ij} = \mu + g_i + e_j + (ge)_{ij} + \varepsilon_{ij} \quad (19)$$

sendo que:

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

μ : é uma constante, geralmente a média;

g_i : é o efeito do i -ésimo genótipo;

e_j : é o efeito do j -ésimo ambiente;

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

ε_{ij} : é erro experimental associado ao i -ésimo genótipo no j -ésimo ambiente, assumido ser independente e $\varepsilon_{ij} \sim N(0, \frac{\sigma^2}{r})$.

Para experimentos dessa natureza, o esquema da análise de variância conjunta é apresentado na Tabela 4.

Tabela 4 - Esquema da análise de variância para experimentos de um mesmo grupo de genótipos avaliado em e locais com r repetições

Fonte de Variação	Graus de liberdade (GL)	Quadrado Médio (QM)
Genótipos (G)	$(g - 1)$	QM_G
Ambientes (E)	$(e - 1)$	QM_E
Interação ($G \times E$)	$(g - 1)(e - 1)$	$QM_{G \times E}$
Erro Médio/r	$ge(r - 1)$	QM_{EM}
Total	$(ger - 1)$	

Como a análise da variância é feita com médias, é necessário que se faça o cálculo do QM_{EM} à parte. Segundo Duarte e Vencovsky (1999), esse valor é obtido fazendo a média poderada dos Quadrados Médios Residuais (QM_{Res}) obtida de todas as ANAVA's individuais de experimentos (Tabela 2), onde os graus de liberdade são os pesos, assim:

$$QM_{EM} = \frac{\sum_j SQ_{Res_j}}{\sum_j GL_{Res_j}} \quad \text{com } j = 1, 2, \dots, e. \quad (20)$$

2.2.3 Análises AMMI

Sendo a interação significativa, o próximo passo é fazer a decomposição da $SQ_{G \times E}$, para descartar um resíduo adicional presente nessa soma de quadrados. Essa decomposição é feita utilizando o fator analítico proposto por Gollob (1968) e Mandel (1969, 1971) e tem a seguinte expressão:

$$(ge)_{ij} = \sum_{k=1}^p \hat{\lambda}_k \alpha_{ik} \gamma_{jk}, \quad (21)$$

em que $\hat{\lambda}_1 \geq \hat{\lambda}_2 \geq \dots \geq \hat{\lambda}_p$, e α_{ik} , γ_{jk} satisfazem o contraste de orto-normalização $\sum_i \alpha_{ik} \alpha'_{ik} = \sum_j \gamma_{jk} \gamma'_{jk} = 0$ para $k \neq k'$ e $\sum_i \alpha_{ik}^2 = \sum_j \gamma_{jk}^2 = 1$.

Antes de aplicar a decomposição, é necessário organizar os dados em uma tabela de dupla entrada (ou matriz de ordem $g \times e$) com as médias dos r blocos (ou repetições) para cada combinação de genótipos e ambientes:

$$\mathbf{Y}_{g \times e} = \begin{pmatrix} Y_{11} & Y_{12} & \dots & Y_{1e} \\ Y_{21} & Y_{22} & \dots & Y_{2e} \\ \dots & \dots & \dots & \dots \\ Y_{g1} & Y_{g2} & \dots & Y_{ge} \end{pmatrix}. \quad (22)$$

O modelo AMMI pressupõe componentes aditivos para os efeitos principais de genótipos e ambientes e componentes multiplicativos para o efeito de interação. Então, a resposta média sobre r repetições ou blocos do i -ésimo genótipo no j -ésimo ambiente é representada por:

$$Y_{ij} = \mu + g_i + e_j + \sum_{k=1}^q \hat{\lambda}_k \alpha_{ik} \gamma_{jk} + \rho_{ij} + \varepsilon_{ij} \quad (23)$$

sendo que:

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

μ : é uma constante, geralmente a média;

g_i : é o efeito do i -ésimo genótipo;

e_j : é o efeito do j -ésimo ambiente;

$\hat{\lambda}_k$: é a raiz quadrada do k -ésimo autovalor da matriz $(\mathbf{GE})(\mathbf{GE})^t$ (ou $(\mathbf{GE})^t(\mathbf{GE})$), com $k = 1, 2, \dots, q$ e onde $q < p$ determina uma aproximação de mínimos quadrados para a matriz $\mathbf{GE}$ pelos q primeiros termos da DVS e $p = \min\{g - 1, e - 1\}$;

α_{ik} : é o i -ésimo elemento do vetor coluna $\boldsymbol{\alpha}_k$ associado a λ_k ;

γ_{jk} : é o j -ésimo elemento do vetor linha $\boldsymbol{\gamma}_k$ associado a λ_k ;

ρ_{ij} : é o resíduo adicional;

ε_{ij} : é erro experimental associado ao i -ésimo genótipo no j -ésimo ambiente, assumido ser independente e $\varepsilon_{ij} \sim N(0, \frac{\sigma^2}{r})$;

Existem várias técnicas para atribuir os graus de liberdade associados à cada parcela da $SQ_{G \times E}$, ou seja, associada a $\hat{\lambda}_k^2$. O método Gollob (1968) destaca-se por ser muito simples e exigir pequenas operações aritméticas. A expressão do método é:

$$GL_{IPCA_k} = g + e - 1 + 2k. \quad (24)$$

A matriz $\mathbf{GE}$ é matriz de interação entre genótipos $\times$ ambiente (matriz de resíduos) obtida do modelo (19), ou seja, cada elemento $(ge)_{ij}$ da matriz $\mathbf{GE}$ é encontrado pela seguinte relação:

$$(ge)_{ij} = Y_{ij} - \bar{Y}_{i.} - \bar{Y}_{.j} + \bar{Y}_{..} \quad (25)$$

onde:

Y_{ij} : é a média das repetições do genótipo i no ambiente j , com $i = 1, 2, \dots, g$ e $j = 1, 2, \dots, e$;

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

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

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

Comparando o modelo (23) com o modelo (19), a interação $(ge)_{ij}$ é representada por $\sum_{k=1}^q \hat{\lambda}_k \alpha_{ik} \gamma_{jk} + \rho_{ij}$, com $q < p$ e $\rho_{ij} = \sum_{k=q+1}^p \hat{\lambda}_k \alpha_{ik} \gamma_{jk}$. E para se encontrar SQ_{Res} pelo método de mínimos quadrados é necessário que se tenham as seguintes restrições ao sistema de equações normais: $\sum_i g_i = \sum_j e_j = \sum_i (ge)_{ij} = \sum_j (ge)_{ij} = 0$.

2.2.4 Correções dos autovalores em matrizes $(\mathbf{GE})(\mathbf{GE})^t$ ou $(\mathbf{GE})^t(\mathbf{GE})$

Para corrigir os autovalores provenientes da matriz $(\mathbf{GE})(\mathbf{GE})^t$ ou $(\mathbf{GE})^t(\mathbf{GE})$, onde $(\mathbf{GE})(\mathbf{GE})^t \sim W_g(e, \Sigma)$, usa-se qualquer um dos três métodos apresentados por Muirhead (1987). Os métodos são:

i) O primeiro método de correção apresentado é :

$$\phi_k^{(1)} = \frac{e \hat{\lambda}_k^2}{\alpha_k^{(1)}} = \frac{e \hat{\lambda}_k^2}{e - g + 1 + 2 \hat{\lambda}_k^2 \sum_{k \neq k^*}^p \frac{1}{\hat{\lambda}_k^2 - \hat{\lambda}_{k^*}^2}} \quad (26)$$

ii) O segundo método de correção apresentado é:

$$\phi_k^{(2)} = \frac{e\hat{\lambda}_k^2}{\alpha_k^{(2)}} = \frac{e\hat{\lambda}_k^2}{e - g - 1 + 2\hat{\lambda}_k^2 \sum_{k \neq k^*}^p \frac{1}{\hat{\lambda}_k^2 - \hat{\lambda}_{k^*}^2}} \quad (27)$$

iii) O terceiro método de correção apresentado é:

$$\begin{aligned} \phi_k^{(3)} &= \frac{e\hat{\lambda}_k^2}{e + g + 1 - 2k} - \frac{ec\hat{\lambda}_k^2 \ln(e\hat{\lambda}_k^2)}{b + \sum_{k=1}^p [\ln(e\hat{\lambda}_k^2)]^2} \\ &= \frac{e\hat{\lambda}_k^2(b + \sum_{k=1}^p [\ln(e\hat{\lambda}_k^2)]^2 - (e + g + 1 - 2k)c \ln(e\hat{\lambda}_k^2))}{(e + g + 1 - 2k)(b + \sum_{k=1}^p [\ln(e\hat{\lambda}_k^2)]^2)} \\ &= \frac{\alpha_k^{(3)}}{d_k} \end{aligned} \quad (28)$$

$$\text{com } b = \frac{5, 8(g-2)^2}{(e+g-1)^2}, \quad c = \frac{6(g-2)}{5(e+g-1)^2},$$

$$\alpha_k^{(3)} = e\hat{\lambda}_k^2(b + \sum_{k=1}^p [\ln(e\hat{\lambda}_k^2)]^2 - (e + g + 1 - 2k)c \ln(e\hat{\lambda}_k^2))$$

$$\text{e } d_k = (e + g + 1 - 2k)(b + \sum_{k=1}^p [\ln(e\hat{\lambda}_k^2)]^2),$$

onde:

g : é número de genótipos do experimento;

e : é número de ambientes do experimento;

$\hat{\lambda}_k^2$: é o k -ésimo autovalor da matriz $(\mathbf{GE})(\mathbf{GE})^t$ (ou $(\mathbf{GE})^t(\mathbf{GE})$), com $k = 1, 2, \dots, p$ e

onde $p = \min\{g-1, e-1\}$;

Os autovalores obtidos com os métodos de correção (26), (27) e (28), nem sempre se apresentam-se na ordem: $\phi_1 \geq \phi_2 \geq \dots \geq \phi_p$. Para colocar os ϕ_i 's na ordem é necessário utilizar a regressão isotônica.

2.2.5 Algoritmo para Isotonização de Stein

Lin e Perlman (1985) apresentam o procedimento de Stein, um algoritmo para modificar autovalores obtidos pelos métodos de correção (26), (27) e (28), ordenando-os de forma crescente.

Para facilitar, listam-se os produtos $e\hat{\lambda}_k^2$ (numerador dos métodos de correção 26 e 27) ou d_k (denominador do método de correção 28) em uma coluna e em outra coluna lista-se o denominador do método de correção (26), $\alpha_k^{(1)}$, ou o denominador do método de correção (27), $\alpha_k^{(2)}$, ou ainda o numerador do método de correção (28), $\alpha_k^{(3)}$, da seguinte maneira:

$$\begin{array}{ccc} e\hat{\lambda}_1^2 \text{ ou } d_1 & \alpha_1^{(1)} \text{ ou } \alpha_1^{(2)} \text{ ou } \alpha_1^{(3)} \\ e\hat{\lambda}_2^2 \text{ ou } d_2 & \alpha_2^{(1)} \text{ ou } \alpha_2^{(2)} \text{ ou } \alpha_2^{(3)} \\ \vdots & \vdots & \vdots \\ e\hat{\lambda}_p^2 \text{ ou } d_p & \alpha_p^{(1)} \text{ ou } \alpha_p^{(2)} \text{ ou } \alpha_p^{(3)} \end{array}$$

Passo 1 Fazendo todos α_k 's positivos:

- a) Inicia-se pelo final da lista e procura-se para cima até que se encontre o primeiro par $(e\hat{\lambda}_k^2, \alpha_k)$ com α_k negativo.
- b) Soma-se este par com o par imediatamente acima dele, substituindo-os pelo par $(e\hat{\lambda}_k^2 + e\hat{\lambda}_{k-1}^2, \alpha_k + \alpha_{k-1})$, para que na lista um par seja menor do que o próximo.
- c) Repete-se (a) e (b), para a nova lista até que todos α_k sejam positivos.

Passo 2 Re-ordenando as razões $\frac{e\hat{\lambda}_k^2}{\alpha_k}$ de forma que estejam em ordem decrescente:

Liste as razões $\frac{e\hat{\lambda}_k^2}{\alpha_k}$ a direita de cada par $(e\hat{\lambda}_k^2, \alpha_k)$ obtidos no passo 1. Um par $(e\hat{\lambda}_k^2, \alpha_k)$ (exceto o par no final da lista) é chamado de par violado se a razão $\frac{e\hat{\lambda}_k^2}{\alpha_k}$ não é maior do que a razão $\frac{e\hat{\lambda}_{k+1}^2}{\alpha_{k+1}}$.

- a) Inicia-se pelo final da lista encontrada no passo 1 e procede-se para cima até o primeiro par violado ser encontrado.
- b) Soma-se este par violado com o par imediatamente acima dele, substitua esses dois pares e suas razões pelo par $(e\hat{\lambda}_k^2 + e\hat{\lambda}_{k+1}^2, \alpha_k + \alpha_{k+1})$ e sua razão $\frac{e\hat{\lambda}_k^2 + e\hat{\lambda}_{k+1}^2}{\alpha_k + \alpha_{k+1}}$, formando uma nova lista.

- c) Re-inicie no par imediatamente após o par trocado em (b) e proceda para cima até o próximo par violado ser encontrado então repita (b).
- d) Repita (c) até todas razões $\frac{e\hat{\lambda}_k^2}{\alpha_k}$ estarem em ordem decrescente.

Passo 3 Cada razão no final da lista foi obtida por bloco acumulado de um ou mais pares consecutivos $(e\hat{\lambda}_k^2, \alpha_k)$ na lista original.

Observação 2.1 *No passo 2 pode ser mais conveniente agrupar qualquer bloco em dois ou mais pares sucessivos na lista cujas razões correspondentes não estão em ordem decrescente.*

2.2.5.1 Número de repetições

A estatística *PRESS* (“Prediction Sum of Squares”), proposta por Allen (1971), consiste em um critério para escolha das variáveis regressoras a serem incluídas no modelo. Especificamente, *PRESS* é a soma de quadrado das diferenças entre um valor observado e um valor predito pelo modelo selecionado. Para os modelos AMMI, Cornélius; Crossa e Seyedsadr (1993) definiram a estatística *PRESS* como:

$$PRESS = \sum_{i=1}^g \sum_{j=1}^e (y_{ij} - \hat{y}_{ij})^2, \quad (29)$$

em que:

y_{ij} : é a média de r repetição no i -ésimo genótipo no j -ésimo ambiente.

$\hat{y}_{ij}$: é a média predita pelo modelo selecionado para o i -ésimo genótipo no j -ésimo ambiente.

Utilizando a estatística *PRESS*, Cornélius; Crossa e Seyedsadr (1993) ajustaram um valor para a medida RMSPD para fazer uma comparação aproximada com outros RMSPDs provenientes de modelos ajustados por validação cruzada. O ajuste é feito por:

$$RMSPD_{(PRESS)} = \sqrt{\frac{PRESS}{ge} + \frac{(r-1)}{r} \frac{QM_{EM}}{r}} \quad (30)$$

em que r é o número de repetição ou o número de blocos em cada ambiente do experimento.

Na maioria dos estudos, existe um grande interesse na comparação do Erro Médio Quadrático do modelo ($QM_{EM}(model)$) selecionado com o Erro Médio Quadrático (QM_{EM}) do experimento. Nachit et al., (1992) utilizaram essa comparação entre o $QM_{EM}(model)$ e o QM_{EM} para encontrar uma aproximação para o número de repetições que falta para o modelo AMMI completo apresentar uma performance igual ao modelo AMMI selecionado, ou seja, indica o número de repetições que se ganha ao analisar os dados com o modelo selecionado. Essa medida é obtida por:

$$R_{AMMI} = \frac{\frac{QM_{EM}}{r}}{QM_{EM}(model)}. \quad (31)$$

Então para fazer uma estimativa do $QM_{EM}(model)$ Piepho (1994) sugere a seguinte expressão:

$$\hat{Q}M_{EM}(model) = (RMSPD)^2 - \frac{QM_{EM}}{r} \quad (32)$$

Assim, este será o procedimento no presente estudo para verificar o ganho em termos de número de repetições ao se fazer a correção dos autovalores nos modelos AMMI.

Todas as análises e a rotina computacional (ver anexo) foram implementadas através do sistema estatístico SAS (SAS Institute Inc., 2003)

2.3 Resultados e Discussão

2.3.1 Experimento 1

2.3.1.1 Análise de variância conjunta

Pela Tabela 5, que corresponde a análise de variância conjunta efetuada com os dados observados, verifica-se, ao nível de 1% de significância que o efeito de blocos é não significativo e que o efeito de genótipos, o efeito de ambientes e o efeito da interação genótipo $\times$ ambiente são significativos e suas somas de quadrados (SQ) correspondem a 1,49%, 72,7% e 9,97%, respectivamente, da soma de quadrados total.

Tabela 5 - Análise de variância conjunta do Experimento 1 com 20 genótipos avaliado em 34 ambientes com 4 blocos

Fonte de Variação	GL	SQ	QM	F	p
Blocos d. ambiente (B d. E)	3	4204574	1401525	3,04	0,0281
Genótipo (G)	19	89066441	4687707	10,16	<0,0001
Ambiente (E)	33	4333925428	131331074	284,51	<0,0001
Interação ($G \times E$)	627	594108485	947541	2,05	<0,0001
Resíduo	2037	940304140	461612		
Total	2719	5961609068			

O resultado de maior interesse nessa tabela é a soma de quadrados da interação genótipo $\times$ ambiente, $SQ_{G \times E} = 594108485$, que pode estar inflacionada devido presença de grande quantidade de ruído na variável resposta.

2.3.1.2 Análise AMMI sem correção dos autovalores

Nesta etapa da análise, é feito um ajuste da interação por Decomposição em Valor Singular (DVS) aplicada a matriz de interação genótipo $\times$ ambiente. Essa matriz tem posto $p = \min(19, 33) = 19$, assim a $SQ_{G \times E}$ pode ser decomposta em 19 componentes ortogonais, que são as somas de quadrados parciais.

Na Tabela 6 é apresentado a análise de cada componente pelo teste F, com os graus de liberdade ajustados pelo método de Gollob (1968). Nota-se, ao nível de 1% de significância que os cinco primeiros componentes são significativos para o modelo, sendo que o primeiro componente retém 30,13% da $SQ_{G \times E}$, o segundo contém 15,38%, 13,29% é retido pelo terceiro componente, 9,08% pelo quarto e 5,76% pelo quinto componente. Esses cinco componentes juntos representam

73,64% da $SQ_{G \times E}$, que é considerada como resposta padrão presente na $SQ_{G \times E}$ com 235 graus de liberdade (37,48% do graus de liberdade da interação) uma vez que a técnica DVS procura maximizar a informação da variabilidade nos primeiros componentes, enquanto que os outros 26,36% da $SQ_{G \times E}$ é devida a ruído presente nos dados.

Tabela 6 - Decomposição de soma de quadrados de interação do Experimento 1 com 20 genótipos avaliado em 34 ambientes com 4 blocos

	GL	$SQ (\lambda_k^2)$	QM	F	valor p
1	51	44764779	877740,76	7,6058707	0
2	49	22850765	466342,14	4,0409858	0
3	47	19744476	420095,24	3,6402435	0
4	45	13493936	299865,25	2,5984168	0
5	43	8556304,5	198983,83	1,7242508	0,0025412
6	41	6905382,4	168423,96	1,459441	0,0305958
7	39	6015703	154248,79	1,3366092	0,0804339
8	37	5804531,5	156879,23	1,3594027	0,0735674
9	35	3882508,1	110928,8	0,9612293	0,5338435
10	33	3550474,4	107590,13	0,9322988	0,5787121
11	31	3216847,4	103769,27	0,8991899	0,6272257
12	29	2512200,4	86627,602	0,7506525	0,8285123
13	27	1903034,8	70482,77	0,610753	0,9421994
14	25	1422974,9	56918,998	0,4932191	0,9833986
15	23	1156819,9	50296,519	0,4358335	0,9909565
16	21	1070682	50984,857	0,4417981	0,9866521
17	19	610242,47	32118,025	0,2783117	0,9991292
18	17	577062,69	33944,864	0,2941418	0,9977349
19	15	488396,51	32559,767	0,2821395	0,9968342

2.3.1.3 Análise AMMI com correção dos autovalores

As estimativas desses componentes ainda podem estar viesadas de acordo com o que foi verificado por Araújo e Dias (2005). Nos modelos AMMI, a retenção de ruído por esses componentes indica que a estimativa do modelo AMMI (como média) não é perfeita (GAUCH JÚNIOR, 1992).

Utilizando as sugestões de Muirhead (1987), obtêm-se os valores dos componentes corrigidos $\phi^{(1)}$, $\phi^{(2)}$ e $\phi^{(3)}$, obtidas pelas eq. (26), (27) e (28), respectivamente, e são apresentados na Tabela 7. Percebe-se que os novos valores dos componentes não satisfazem algumas condições que deveriam ser respeitadas. A ordem decrescente dos valores não é verificada para nenhum dos métodos de correção. Os valores de $\phi^{(1)}$ e $\phi^{(2)}$ apresentam outro problema que é fato de assumirem valores negativos. Para superar esses problemas utiliza-se da regressão isotônica, mais propriamente o algoritmo de Steim apresentado por Lin e Perlman (1985). Os valores $\phi^{(1)*}$, $\phi^{(2)*}$ e

$\phi^{(3)*}$, também são apresentados na Tabela 7 e se referem a regressão isotônica de $\phi^{(1)}$, $\phi^{(2)}$ e $\phi^{(3)}$, respectivamente.

Tabela 7 - Correção dos autovalores da matriz $(GE)(GE)^t$ do experimento 1 e os autovalores ajustados pela regressão isotônica

	λ^2	$\phi^{(1)}$	$\phi^{(1)*}$	$\phi^{(2)}$	$\phi^{(2)*}$	$\phi^{(3)}$	$\phi^{(3)*}$
1	44764779	26268136	26268136	27207272	27207272	6286366,5	6286366,5
2	22850765	11426689	12962548	11772993	13443869	4147974	4168439,5
3	19744476	15350388	12962548	16086043	13443869	4189740,4	4168439,5
4	13493936	9286514,5	9286514,5	9678314,9	9678314,9	3383543,8	3383543,8
5	8556304,5	4787919,2	4957793,3	4950884,2	5179251,3	2512930,8	2512930,8
6	6905382,4	3759430,2	4957793,3	3883808,1	5179251,3	2305817,1	2330966,4
7	6015703	2391437,2	4957793,3	2448698,2	5179251,3	2260432,4	2330966,4
8	5804531,5	-9118598	4957793,3	-8347242	5179251,3	2432846,1	2330966,4
9	3882508,1	2142554,1	3602683,1	2214438,4	3831408,4	1847233,7	1876806,4
10	3550474,4	3884208,8	3602683,1	4151360,3	3831408,4	1883214,1	1876806,4
11	3216847,4	13991748	3602683,1	18802428	3831408,4	1903167,5	1876806,4
12	2512200,4	4050462,3	3602683,1	4474868,4	3831408,4	1667989,7	1667989,7
13	1903034,8	2732130,3	2732130,3	2984144,9	2984144,9	1420000,3	1420000,3
14	1422974,9	1710600,2	2419297,4	1840767,3	2739715,8	1195173,1	1195173,1
15	1156819,9	1125535,6	2419297,4	1193863,7	2739715,8	1093842,4	1116427,8
16	1070682	-3052528	2419297,4	-2614122	2739715,8	1140977,1	1116427,8
17	610242,47	402691,41	2349044,5	418953,91	2739715,8	744352,27	772206,06
18	577062,69	-1219441	2349044,5	-1084618	2739715,8	802991,83	772206,06
19	488396,51	-1485228	2349044,5	-1259859	2739715,8	784362,66	772206,06
Total	148527121		112758807		119238403		41975675
$\frac{\sum_{k=1}^{19} \phi_k^{(1)*}}{\sum_{k=1}^{19} \lambda_k^2}$			0,7591		0,8028		0,2826

Verifica-se ainda que $\sum_{k=1}^{19} \phi_k^{(1)*}$ é equivalente a 75,91% da $SQ_{G \times E}$ e o restante da $SQ_{G \times E}$ (24,09%) representa ruído detectado pela correção que estava presente na interação. O $\sum_{k=1}^{19} \phi_k^{(2)*}$ é equivalente a 80,28% e 19,72% representa o ruído excluído da $SQ_{G \times E}$ pela correção e $\sum_{k=1}^{19} \phi_k^{(3)*}$ representa 28,26% da $SQ_{G \times E}$, sendo que a correção considerou 71,74% da interação genótipos $\times$ ambientes como sendo ruído.

Na Tabela 8 apresenta-se a análise de cada componente pelo teste F, com os graus de liberdade ajustados pelo método de Gollob (1968), para os autovalores corrigidos pelo método 1 (eq. (26)). Considerando um nível de significância de 1%, tem-se que os quatros primeiros componentes são significativos para o modelo, sendo que o primeiro, o segundo, o terceiro e o quarto componente retém, respectivamente, 17,68%, 8,72%, 8,72% e 6,25% da $SQ_{G \times E}$. A união dos quatros componentes representam 41,37% da $SQ_{G \times E}$ com 192 graus de liberdade, correspondente a 30,62 % dos graus de liberdade da interação, enquanto que os outros 58,61% da $SQ_{G \times E}$ são ruídos presente nos dados descartados pelo teste F e pelo método de correção 1.

Tabela 8 - Análise do Teste F para os componentes corrigidos do Experimento 1 pelo método 1

	GL	$SQ (\phi_k^{(1)*})$	QM	F	valor p
1	51	26268136	515061,48	4,4631527	0
2	49	12962548	264541,79	2,2923291	0
3	47	12962548	275798,89	2,389875	0
4	45	9286514,5	206366,99	1,788228	0,0010826
5	43	4957793,3	115297,52	0,9990854	0,4738763
6	41	4957793,3	120921,79	1,0478213	0,3889046
7	39	4957793,3	127122,9	1,1015557	0,3076563
8	37	4957793,3	133994,41	1,1610993	0,2336865
9	35	3602683,1	102933,8	0,8919504	0,6506715
10	33	3602683,1	109172,21	0,946008	0,5560882
11	31	3602683,1	116215,58	1,0070407	0,456096
12	29	3602683,1	124230,45	1,0764918	0,3566964
13	27	2732130,3	101190,01	0,8768399	0,6478753
14	25	2419297,4	96771,894	0,8385557	0,6940435
15	23	2419297,4	105186,84	0,9114736	0,5832825
16	21	2419297,4	115204,64	0,9982806	0,4617659
17	19	2349044,5	123633,92	1,0713227	0,3744749
18	17	2349044,5	138179,09	1,1973607	0,25777
19	15	2349044,5	156602,97	1,3570088	0,1600472

A Tabela 9 apresenta a análise de cada componente pelo teste F , com os graus de liberdade ajustados pelo método de Gollob (1968), para os autovalores corrigidos pelo método 2 (eq. (27)). Considerando um nível de significância de 1%, tem-se que os quatros primeiros componentes são significativos para o modelo, sendo que o primeiro, o segundo, o terceiro e o quarto componente retém, respectivamente, 18,32%, 9,05%, 9,05% e 6,51% da $SQ_{G \times E}$. A união dos quatros componentes representam 42,93% da $SQ_{G \times E}$ com 192 graus de liberdade, que representam 30,62 % do graus de liberdade da interação, enquanto que os outros 58,61% da $SQ_{G \times E}$ são ruídos presente nos dados, descartados pelo teste F e pelo método de correção 2.

A Tabela 10 apresenta a análise de cada componente pelo teste F , com os graus de liberdade ajustados pelo método de Gollob (1968), para os autovalores corrigidos pelo método 3 (eq. (28)). Considerando um nível de significância de 1%, tem-se que todos componentes são não significativos para o modelo, assim o método 3 e o teste F considera que toda interação genótipo $\times$ ambiente é composta por ruídos, logo modelo aditivo é o melhor para analisar os dados.

Tabela 9 - Análise do Teste F para os componentes corrigidos do Experimento 1 pelo método 2

	GL	SQ ($\phi_k^{(2)*}$)	QM	F	valor p
1	51	27207272	533475,93	4,6227191	0
2	49	13443869	274364,67	2,377447	0
3	47	13443869	286039,76	2,478615	0
4	45	9678314,9	215073,66	1,8636738	0,0004788
5	43	5179251,3	120447,71	1,0437133	0,3946925
6	41	5179251,3	126323,2	1,0946261	0,3148536
7	39	5179251,3	132801,32	1,1507608	0,2417099
8	37	5179251,3	139979,77	1,2129641	0,1779072
9	35	3831408,4	109468,81	0,948578	0,5552697
10	33	3831408,4	116103,28	1,0060676	0,4583626
11	31	3831408,4	123593,82	1,0709752	0,3619375
12	29	3831408,4	132117,53	1,1448356	0,2717196
13	27	2984144,9	110523,88	0,9577206	0,5267683
14	25	2739715,8	109588,63	0,9496163	0,5346603
15	23	2739715,8	119118,08	1,0321917	0,4191896
16	21	2739715,8	130462,66	1,1304956	0,3071713
17	19	2739715,8	144195,57	1,2494952	0,207855
18	17	2739715,8	161159,75	1,3964946	0,128109
19	15	2739715,8	182647,72	1,5826939	0,0707614

Tabela 10 - Análise do Teste F para os componentes corrigidos do Experimento 1 pelo método 3

	GL	SQ ($\phi_k^{(3)*}$)	QM	F	valor p
1	51	6286366,5	123262,09	1,0681007	0,3463179
2	49	4147974	84652,53	0,733538	0,9158261
3	47	4189740,4	89143,413	0,7724528	0,8687266
4	45	3383543,8	75189,863	0,6515413	0,9648924
5	43	2512930,8	58440,251	0,5064012	0,9969411
6	41	2305817,1	56239,442	0,4873306	0,9975462
7	39	2260432,4	57959,804	0,502238	0,9957852
8	37	2432846,1	65752,598	0,5697648	0,9829754
9	35	1847233,7	52778,106	0,4573371	0,9974336
10	33	1883214,1	57067,094	0,4945024	0,9931069
11	31	1903167,5	61392,5	0,5319833	0,9842332
12	29	1667989,7	57516,886	0,4984	0,988551
13	27	1420000,3	52592,603	0,4557297	0,9928415
14	25	1195173,1	47806,924	0,4142604	0,9954682
15	23	1093842,4	47558,366	0,4121066	0,9939441
16	21	1140977,1	54332,241	0,4708042	0,9800704
17	19	744352,27	39176,435	0,3394748	0,9965599
18	17	802991,83	47234,813	0,409303	0,9839106
19	15	784362,66	52290,844	0,4531149	0,9627846

2.3.1.4 Número de Repetições

De acordo com interpretação dada por Nachit, et al., (1992), a medida R_{AMMI} indica uma aproximação para o número de repetições que falta para o modelo AMMI completo apresentar

uma performance igual ao modelo AMMI selecionado, ou seja, a medida R_{AMMI} indica o número de repetições que se ganha ao analisar os dados com o modelo selecionado. Assim, corrigindo os componentes pelos métodos 1 e 2 têm-se um ganho de aproximadamente 3 repetições, enquanto que para o método 3 tem-se um ganho aproximado de uma repetição. Já em relação medida $RMSPD_{PRESS}$ o menor valor foi obtido pelo método de correção 2 e o maior foi obtido pelo método 3, vide Tabela 11.

Tabela 11 - $RMSPD_{PRESS}$ e R_{AMMI} para o melhor modelo AMMI selecionado para os dados do Experimento 1 após a correção dos autovalores pelos métodos 1, 2 e 3

	$RMSPD_{PRESS}$	R_{AMMI}
método 1	404,88585	2,3779988
método 2	403,7749	2,4228505
método 3	552,24499	0,6087575

2.3.2 Experimento 2

2.3.2.1 Análise de variância conjunta

Pelo quadro da análise de variância conjunta do experimento 2 (Tabela 12), nota-se, ao nível de 1% de significância, que os efeitos de genótipos, de ambientes e a interação genótipo $\times$ ambiente são significativos, sendo que cada efeito representa 2,03%, 55,60% e 11,96% , respectivamente da soma de quadrados total, enquanto que o efeito de blocos dentro de ambientes é não significativo.

Tabela 12 - Análise de variância conjunta do Experimento 2 com 16 genótipos avaliado em 24 ambientes com 4 blocos

Fonte de Variação	GL	SQ	QM	F	p
Blocos d. ambiente (B d. E)	3	11478022	3826007	0,96	0,4088
Genótipo (G)	15	304645583	20309706	5,12	<0,0001
Ambiente (E)	23	8355949394	363302148	91,57	<0,0001
Interação ($G \times E$)	345	1797966686	5211498	1,31	0,0006
Resíduo	1149	4558409355	3967284		
Total	1535	15028449039			

Neste experimento, tem-se o interesse em decompor os 11,96% da SQ_{total} , em ruído e padrão, sendo que este valor da $SQ_{G \times E}$ pode estar inflacionado de ruídos.

2.3.2.2 Análise AMMI sem correção dos autovalores

Ao analisar, por Decomposição em Valor Singular aplicado a matriz de interação genótipo $\times$ ambiente que tem posto $p = \min(15, 23) = 15$, tem-se que a $SQ_{G \times E}$ é decomposta em 15 componentes ortogonais.

Tabela 13 - Decomposição de soma de quadrados de interação do Experimento 2 com 16 genótipos avaliados em 24 ambientes com 4 blocos

	GL	SQ (λ_k^2)	QM	F	p
1	37	210053838	5677130,7	5,7239469	0
2	35	54411361	1554610,3	1,5674303	0,0195223
3	33	39073038	1184031,5	1,1937955	0,2098402
4	31	36036876	1162479,9	1,1720662	0,238002
5	29	27876854	961270,82	0,9691979	0,5129259
6	27	19567090	724707,03	0,7306833	0,8401437
7	25	16514580	660583,19	0,6660306	0,8925159
8	23	14536619	632026,89	0,6372389	0,9051798
9	21	9063067,4	431574,64	0,4351336	0,9877539
10	19	7926903,7	417205,46	0,4206459	0,9864339
11	17	6026830,1	354519,42	0,3574429	0,9924881
12	15	3795606,6	253040,44	0,2551271	0,99821
13	13	2254479,1	173421,47	0,1748516	0,9995274
14	11	1538305,9	139845,99	0,1409992	0,9995454
15	9	816224,39	90691,599	0,0914395	0,9997458

Utilizando o método de Gollob (1968) para atribuir graus de liberdade a cada componente, é realizado uma análise pelo teste F e os resultados são apresentados na Tabela 13. Ao nível de 1% de significância, nota-se que antes de corrigir os autovalores têm-se que um componente é significativo para ser retido no modelo selecionado, sendo que este componente corresponde a 46,73% (parte da interação considerada como resposta padrão pelo teste F) da soma de quadrados da interação genótipo $\times$ ambiente, utilizando 37 graus de liberdade (10,72%) do total de graus de liberdade da interação.

2.3.2.3 Análise AMMI com correção dos autovalores

Utilizando as sugestões de Muirhead (1987), obtêm-se os valores dos componentes corrigidos $\phi^{(1)}$, $\phi^{(2)}$ e $\phi^{(3)}$, obtidos pelas eq. (26), (27) e (28), respectivamente, e as correções são apresentados na Tabela 14. Os novos valores dos componentes não satisfaz uma condição que deveria ser respeitada, a saber, a ordem decrescente dos valores que não é verificada para nenhum

dos métodos de correção. Utilizado o algoritmo de Steim apresentado por Lin e Perlman (1985), que utiliza regressão isotônica, pode-se superar esse problema, tornado os valores das componentes ordenados. Os valores obtidos pela regressão isotônica $\phi^{(1)*}$, $\phi^{(2)*}$ e $\phi^{(3)*}$ de $\phi^{(1)}$, $\phi^{(2)}$ e $\phi^{(3)}$, respectivamente, também são apresentados na Tabela 14.

Tabela 14 - Correção dos autovalores da matriz $(GE)(GE)^t$ do experimento 2 e os autovalores ajustados pela regressão isotônica

	λ^2	$\phi^{(1)}$	$\phi^{(1)*}$	$\phi^{(2)}$	$\phi^{(2)*}$	$\phi^{(3)}$	$\phi^{(3)*}$
1	210053838	126952690	126952690	133685794	133685794	27487136	27487136
2	54411361	26403637	26403637	27516352	27516352	10524103	10524103
3	39073038	14941493	22510070	15433299	23693543	9286283,6	9693575,6
4	36036876	49936383	22510070	56455605	23693543	10125552	9693575,6
5	27876854	24280892	22510070	26181225	23693543	9295999,9	9295999,9
6	19567090	12307410	16857770	12988192	18388809	7719914,2	7719914,2
7	16514580	13944432	16857770	14999888	18388809	7588725,2	7668065,8
8	14536619	63619786	16857770	100142890	18388809	7753753,7	7668065,8
9	9063067,4	6919506,3	10653544	7389663,8	11896863	5683626,9	5730535,2
10	7926903,7	27815009	10653544	39309580	11896863	5781910,8	5730535,2
11	6026830,1	20088547	10653544	27814440	11896863	5156842,5	5156842,5
12	3795606,6	8743829,6	8743829,6	10821204	10821204	3849816,8	3849816,8
13	2254479,1	4759231,2	5543222,1	5775188,7	7328293,9	2736521,4	2736521,4
14	1538305,9	7307390,4	5543222,1	12095465	7328293,9	2258631,7	2258631,7
15	816224,39	5130990,4	5130990,4	10776090	7328293,9	1483625	1483625
Total	449491672		328381743		355945877		116696944
$\frac{\sum_{k=1}^{15} \phi_k^{(1)*}}{\sum_{k=1}^{15} \lambda_k^2}$			0,7305		0,7918		0,2596

Verifica-se ainda na Tabela 14 que $\sum_{k=1}^{15} \phi_k^{(1)*}$ é equivale a 73,05% da $SQ_{G \times E}$ e o restante da $SQ_{G \times E}$ (26,95%) representa ruído detectado pela correção que estava presente na interação. A $\sum_{k=1}^{15} \phi_k^{(2)*}$ é equivalente a 79,18% do total da $SQ_{G \times E}$ e 20,82% representa o ruído excluído da $SQ_{G \times E}$ pela correção e $\sum_{k=1}^{15} \phi_k^{(3)*}$ representa 25,96% da $SQ_{G \times E}$, sendo que a correção considerou 74,04% da interação genótipos $\times$ ambientes como sendo ruído.

Na Tabela 15 é apresentada a análise de cada componente pelo teste F para os autovalores corrigidos pelo método 1 (eq. (26)), com os graus de liberdade ajustados pelo método de Gollob (1968). Considerando um nível de significância de 1%, tem-se que somente o primeiro componente é significativo para ser retido no modelo, sendo que este componente retém 28,24% e descarta 71,76% da $SQ_{G \times E}$, utilizando apenas 37 graus de liberdade (10,72%) de um total de 345 graus de liberdade.

Tabela 15 - Análise do Teste F para os componentes do Experimento 2 corrigidos pelo método 1

	GL	$SQ(\phi_k^{(1)*})$	QM	F	valor p
1	37	126952690	3431153,8	3,4594486	0
2	35	26403637	754389,64	0,7606107	0,8420432
3	33	22510070	682123,35	0,6877484	0,9089009
4	31	22510070	726131,3	0,7321193	0,8578178
5	29	22510070	776209,33	0,7826103	0,788417
6	27	16857770	624361,84	0,6295106	0,929751
7	25	16857770	674310,79	0,6798714	0,8803146
8	23	16857770	732946,51	0,7389907	0,8079728
9	21	10653544	507311,6	0,5114951	0,9669073
10	19	10653544	560712,82	0,5653367	0,9313197
11	17	10653544	626679,03	0,6318469	0,8687082
12	15	8743829,6	582921,97	0,587729	0,8861427
13	13	5543222,1	426401,7	0,429918	0,9594464
14	11	5543222,1	503929,28	0,5080849	0,898845
15	9	5130990,4	570110,04	0,5748114	0,818578

Pela Tabela 16, que apresenta a análise de cada componente pelo teste F , com os graus de liberdade ajustados pelo método de Gollob (1968), para os autovalores corrigidos pelo método 2 (eq. (27)). Para um nível de significância de 1%, tem-se que somente o primeiro componente é significativo para ser retido no modelo, sendo que este componente retém 29,74% e descarta 70,26% da $SQ_{G \times E}$, utilizando apenas 37 graus de liberdade (10,72%) de um total de 345 graus de liberdade.

Tabela 16 - Análise do Teste F para os componentes do Experimento 2 corrigidos pelo método 2

	GL	$SQ(\phi_k^{(1)*})$	QM	F	valor p
1	37	133685794	3613129,6	3,642925	0
2	35	27516352	786181,49	0,7926647	0,8011836
3	33	23693543	717986,16	0,723907	0,8749387
4	31	23693543	764307,85	0,7706107	0,8129389
5	29	23693543	817018,74	0,8237562	0,7328325
6	27	18388809	681066,99	0,6866834	0,8843913
7	25	18388809	735552,35	0,741618	0,8169874
8	23	18388809	799513,42	0,8061066	0,7264456
9	21	11896863	566517,28	0,571189	0,9387716
10	19	11896863	626150,67	0,6313142	0,8847319
11	17	11896863	699815,46	0,7055864	0,7995005
12	15	10821204	721413,62	0,7273627	0,7582254
13	13	7328293,9	563714,91	0,5683636	0,8803637
14	11	7328293,9	666208,53	0,6717024	0,7663491
15	9	7328293,9	814254,88	0,8209696	0,5968456

A Tabela 17 mostra a análise de cada componente pelo teste F para os autovalores corrigidos pelo método 3 (eq. (28)), com os graus de liberdade ajustados pelo método de Gollob

(1968). Considerando um nível de significância de 1%, tem-se que nenhum componente é significativo para o modelo, sendo assim esta correção considera que toda interação genótipo $\times$ ambiente é composta por ruídos, e portanto tem-se um modelo aditivo para análise dos dados.

Tabela 17 - Análise do Teste F para os componentes do Experimento 2 corrigidos pelo método 3

	GL	SQ ($\phi_k^{(3)*}$)	QM	F	valor p
1	37	27487136	742895,57	0,7490218	0,8631103
2	35	10524103	300688,65	0,3031683	0,999975
3	33	9286283,6	281402,53	0,2837231	0,9999811
4	31	10125552	326630,71	0,3293242	0,999828
5	29	9295999,9	320551,72	0,3231951	0,9997748
6	27	7719914,2	285922,75	0,2882806	0,9998804
7	25	7588725,2	303549,01	0,3060522	0,9996405
8	23	7753753,7	337119,73	0,3398998	0,9985824
9	21	5683626,9	270648,9	0,2728808	0,9995821
10	19	5781910,8	304311,1	0,3068206	0,9982444
11	17	5156842,5	303343,68	0,3058452	0,997073
12	15	3849816,8	256654,45	0,2587709	0,9980557
13	13	2736521,4	210501,65	0,2122375	0,9986473
14	11	2258631,7	205330,15	0,2070234	0,9972239
15	9	1483625	164847,23	0,1662066	0,9971509

2.3.2.4 Número de Repetições

Tabela 18 - $RMSPD_{PRESS}$ e R_{AMMI} para o melhor modelo AMMI selecionado para os dados do Experimento 2 após a correção dos autovalores pelos métodos 1, 2 e 3

	$RMSPD_{PRESS}$	R_{AMMI}
método 1	1180,8903	2,4630443
método 2	1178,8864	2,4923119
método 3	1383,6246	1,0750329

Na Tabela 18 são apresentados os valores das medidas $RMSPD_{PRESS}$ e R_{AMMI} e têm-se que corrigindo os componentes pelos métodos 1 e 2 têm-se um ganho de aproximadamente 3 repetições, enquanto que para o método 3 tem-se um ganho aproximado para duas repetição. Já em relação medida $RMSPD_{PRESS}$ o menor valor foi obtido pelo método de correção 2 (1178,8864) e o maior foi obtido pelo método 3 (1383,6246).

2.3.3 Experimento 3

2.3.3.1 Análise de variância conjunta

Pela Tabela 19, que corresponde à parte da análise de variância conjunta efetuada com os dados observados. Verifica-se, ao nível de 1% de significância que o efeito de blocos dentro de ambientes é não significativo e o efeito de genótipos, o efeito de ambientes e o efeito da interação genótipo $\times$ ambiente são significativos e suas somas de quadrados (SQ) corresponde a 4,62%, 57,29,7% e 14,46%, respectivamente, da soma de quadrados total.

Tabela 19 - Análise de variância conjunta do Experimento 3 com 9 genótipos avaliado em 20 ambientes com 4 blocos

Fonte de Variação	GL	SQ	QM	F	valor p
Blocos d. ambiente (B d. E)	3	8265754,2	2755251,4	3,70	0,0117
Genótipo (G)	8	79828574,7	9978571,8	13,40	<0,0001
Ambiente (E)	19	989593771,8	52083882,7	69,94	<0,0001
Interação ($G \times E$)	152	249704161,7	1642790,5	2,21	<0,0001
Resíduo	537	399913799	744718		
Total	719	1727306061			

O resultado de maior interesse nessa tabela é a soma de quadrados da interação genótipo $\times$ ambiente, $SQ_{G \times E} = 249704161,7$, que pode estar inflacionada, devido a grande presença de ruído na resposta.

2.3.3.2 Análise AMMI sem correção dos autovalores

Nesta etapa da análise, é feito um ajuste da interação por Decomposição em Valor Singular (DVS) aplicada a matriz de interação genótipo $\times$ ambiente (GE). Essa matriz tem posto $p = \min(8, 19) = 8$, assim a $SQ_{G \times E}$ pode ser decomposta em 8 eixos ortogonais, que são as somas de quadrados parciais.

Na Tabela 20 é apresentada a análise de cada componente pelo teste F, com os graus de liberdade ajustados pelo método de Gollob (1968). Nota-se, ao nível de 1% de significância, que os dois primeiros eixos são significativos para o modelo, sendo que o primeiro componente retém 56,20% da $SQ_{G \times E}$ e o segundo componente contém 15,10% da $SQ_{G \times E}$. Esses dois componentes juntos representam 71,30% da $SQ_{G \times E}$, que é resposta padrão presente na $SQ_{G \times E}$ com 50 graus de liberdade (32,89% do graus de liberdade da interação), enquanto que os outros 28,70% da $SQ_{G \times E}$ são ruídos presente nos dados, sem fazer a correção dos autovalores da matriz $(GE)(GE)'$.

Tabela 20 - Decomposição de soma de quadrados de interação do Experimento 3 com 9 genótipos avaliado em 20 ambientes com 4 blocos

	GL	SQ (λ_k^2)	QM	F	valor p
1	26	35080855	1349263,6	7,2471075	0
2	24	9427908,4	392829,52	2,1099492	0,001701
3	22	6514643,1	296120,14	1,5905079	0,0434142
4	20	5383283,3	269164,17	1,4457231	0,0952784
5	18	3092926	171829,22	0,9229218	0,5505275
6	16	1725058,9	107816,18	0,5790977	0,9002883
7	14	628267,72	44876,265	0,2410375	0,9981205
8	12	573098,44	47758,204	0,2565168	0,9948124

2.3.3.3 Análise AMMI com correção dos autovalores

A estimativas desses componentes ainda podem estar viesadas de acordo com o que foi verificado por Araújo e Dias (2005). No caso dos modelos AMMI a retenção de ruído por esses componentes indica que a estimativa do modelo AMMI (como média) não é perfeita (GAUCH JÚNIOR, 1992).

Tabela 21 - Correção dos autovalores da matriz $(GE)(GE)^t$ do experimento 3 e os autovalores ajustados pela regressão isotônica

	λ^2	$\phi^{(1)}$	$\phi^{(1)*}$	$\phi^{(2)}$	$\phi^{(2)*}$	$\phi^{(3)}$	$\phi^{(3)*}$
1	35080855	25129385	25129385	27068367	27068367	11440387	11440387
2	9427908,4	5874835,4	5874835,4	6265243	6344715,5	3828678,1	3828678,1
3	6514643,1	4413339,3	5785800,6	4734046,9	6344715,5	3109093,4	3109093,4
4	5383283,3	9277117,1	5785800,6	11208741	6344715,5	2996549,3	2996549,3
5	3092926	3912999	3912999	4479752,5	4479752,5	2035485,1	2035485,1
6	1725058,9	2574023,9	2574023,9	3025464,6	3025464,6	1346145,5	1346145,5
7	628267,72	387060,41	1110123,9	412471,79	1361798,1	589619,82	613684,55
8	573098,44	-1059358	1110123,9	-894088,4	1361798,1	641187,09	613684,55
Total	62426040		51283092		56331327		25983707
$\frac{\sum_{k=1}^8 \phi_k^{(i)*}}{\sum_{k=1}^8 \lambda_k^2}$			0,8215		0,9024		0,4162

Utilizando as sugestões de Muirhead (1987), obtêm-se os valores dos componentes corrigidos $\phi^{(1)}$, $\phi^{(2)}$ e $\phi^{(3)}$, obtidos pelos métodos de correções 1, 2 e 3 (eq. (26), (27) e (28)), respectivamente, e são apresentados na Tabela 21. Percebe-se que os novos valores dos componentes não satisfazem algumas condições que deveriam ser respeitadas. A ordem decrescente dos valores não é verificada para nenhum dos métodos de correção. Os valores de $\phi^{(1)}$ e $\phi^{(2)}$ apresentam outros problema assumindo valores negativos. Para superar esses problemas utiliza-se da regressão isotônica, mais propriamente utilizado o algoritmo de Steim apresentado por Lin e Perlman (1985). Os valores $\phi^{(1)*}$, $\phi^{(2)*}$ e $\phi^{(3)*}$ também são apresentados na Tabela 21 e se referem a regressão

isotônica de $\phi^{(1)}$, $\phi^{(2)}$ e $\phi^{(3)}$, respectivamente.

Verifica-se ainda que $\sum_{k=1}^8 \phi_k^{(1)*}$ é equivalente a 82,15% da $SQ_{G \times E}$ e o restante da $SQ_{G \times E}$ (17,85%) representa ruído detectado pela correção que estava presente na interação. O $\sum_{k=1}^8 \phi_k^{(2)*}$ é equivalente a 90,24% e 9,76% representa o ruído excluído da $SQ_{G \times E}$ pela correção e $\sum_{k=1}^{19} \phi_k^{(3)*}$ representa 41,62% da $SQ_{G \times E}$, sendo que a correção considerou 58,38% da interação genótipos $\times$ ambientes como sendo ruído.

Tabela 22 - Análise do Teste F para os componentes do Experimento 3 corrigidas pelo método 1

	GL	$SQ (\phi_k^{(1)*})$	QM	F	valor p
1	26	25129385	966514,81	5,1913033	0
2	24	5874835,4	244784,81	1,3147778	0,1457344
3	22	5785800,6	262990,94	1,4125657	0,1009534
4	20	5785800,6	289290,03	1,5538223	0,0590988
5	18	3912999	217388,83	1,1676297	0,2833828
6	16	2574023,9	160876,49	0,864093	0,6115569
7	14	1110123,9	79294,562	0,4259036	0,9665967
8	12	1110123,9	92510,322	0,4968875	0,9168447

Na Tabela 22, apresenta-se a análise de cada componente pelo teste F , com os graus de liberdade ajustados pelo método de Gollob (1968), para os autovalores corrigidos pelo método 1 (eq. (26)). Considerando um nível de significância de 1%, tem-se que somente o primeiro componente é significativo para o modelo, sendo que este primeiro componente retém 40,25% da $SQ_{G \times E}$ com 26 graus de liberdade, que representam 17,10% do total de graus de liberdade da interação. O restante da $SQ_{G \times E}$ (59,75%) é descartado pela correção e pelo teste F , sendo considerados como ruídos presente nos dados.

A Tabela 23 apresenta a análise de cada componente pelo teste F , com os graus de liberdade ajustados pelo método de Gollob (1968), para os autovalores corrigidos pelo método 2 (eq. (27)). Considerando um nível de significância de 1%, tem-se que somente o primeiro componente é significativo para o modelo, sendo que este primeiro componente retém 43,36% da $SQ_{G \times E}$ com 26 graus de liberdade, que representam 17,10% do grau de liberdade da interação. O restante da $SQ_{G \times E}$ (56,64%) é descartada pela correção e pelo teste F , sendo considerados como ruídos presente nos dados.

Tabela 23 - Análise do Teste F para os componentes do Experimento 3 corrigidas pelo método 2

	GL	SQ ($\phi_k^{(2)*}$)	QM	F	valor p
1	26	27068367	1041091	5,5918639	0
2	24	6344715,5	264363,15	1,4199361	0,0899288
3	22	6344715,5	288396,16	1,5490212	0,0532884
4	20	6344715,5	317235,78	1,7039233	0,0290366
5	18	4479752,5	248875,14	1,3367476	0,1587889
6	16	3025464,6	189091,54	1,0156404	0,4379339
7	14	1361798,1	97271,293	0,5224594	0,9205518
8	12	1361798,1	113483,17	0,609536	0,8348821

Na Tabela 24 é feita uma análise de cada componente pelo teste F , com os graus de liberdade ajustados pelo método de Gollob (1968), para os autovalores corrigidos pelo método 3 (eq. (28)). Considerando um nível de significância de 1%, tem-se que somente o primeiro componente é significativo para o modelo, sendo que este primeiro componente retém 43,36% da $SQ_{G \times E}$ com 26 graus de liberdade, que representam 17,10% do graus de liberdade da interação. O restante da $SQ_{G \times E}$ (56,64%) é descartado pela correção e pelo teste F , sendo considerados como ruídos presente nos dados.

Tabela 24 - Análise do Teste F para os componentes do Experimento 3 corrigidas pelo método 3

	GL	SQ ($\phi_k^{(3)*}$)	QM	F	valor p
1	26	11440387	440014,88	2,3633892	0,000198
2	24	3828678,1	159528,25	0,8568514	0,6627256
3	22	3109093,4	141322,43	0,759065	0,7770384
4	20	2996549,3	149827,46	0,8047469	0,7089841
5	18	2035485,1	113082,51	0,6073839	0,8950522
6	16	1346145,5	84134,093	0,4518975	0,967581
7	14	589619,82	42115,702	0,2262101	0,9986811
8	12	641187,09	53432,257	0,2869931	0,991278

2.3.3.4 Número de Repetições

São apresentados na Tabela 25 os valores das medidas $RMSPD_{PRESS}$ e R_{AMMI} e vê-se que corrigindo os componentes pelo método 2 têm-se um ganho de aproximadamente 3 repetições, enquanto que para os métodos 1 e 3 tem-se um ganho aproximado de 2 repetições. Já em relação à medida $RMSPD_{PRESS}$ o menor valor foi obtido pelo método de correção 2 (514,53478) e o maior foi obtido pelo método 3 (572,19753).

Tabela 25 - $RMSPD_{PRESS}$ e R_{AMMI} para o melhor modelo AMMI selecionado para os dados do Experimento 3 após a correção dos autovalores pelos métodos 1, 2 e 3

	$RMSPD_{PRESS}$	R_{AMMI}
Método 1	544,19926	1,692954
Método 2	514,53478	2,3697094
Método 3	572,19753	1,3182686

2.3.4 Experimento 4

2.3.4.1 Análise de variância conjunta

Tabela 26 - Análise de variância conjunta do Experimento 4 com 8 genótipos avaliado em 59 ambientes com 4 blocos

Fonte de Variação	GL	SQ	QM	F	valor p
Blocos d. ambiente (B d. E)	3	11145643	3715214	6,26	0,0003
Genótipo (G)	7	82527757	11789680	19,87	<0,0001
Ambiente (E)	58	3787154892	65295774	110,03	<0,0001
Interação ($G \times E$)	406	263183718	648236	1,09	0,1288
Resíduo	1413	838537156	593445		
Total	1887	4982549167			

Pelo quadro da análise de variância conjunta do experimento 4 (Tabela 26), nota-se, ao nível de 1% de significância que os efeitos de bloco dentro ambiente, de genótipo e de ambiente são significativos, sendo que cada efeito representa 0,22%, 1,65% e 76,01% respectivamente da soma de quadrados total, já a interação genótipo $\times$ ambiente, é não significativa, mas mesmo assim é de interesse decompor essa soma de quadrados $SQ_{G \times E}$, para verificar se os resultados concordam com a análise da variância.

2.3.4.2 Análise AMMI sem correção dos autovalores

Tabela 27 - Decomposição de soma de quadrados de interação do Experimento 4 com 8 genótipos avaliados em 59 ambientes com 4 blocos

	GL	SQ (λ_k^2)	QM	F	valor p
1	64	16382511	255976,73	1,7253624	0,0004221
2	62	12011970	193741,45	1,3058773	0,0582035
3	60	10890890	181514,84	1,2234662	0,1205794
4	58	9447805,9	162893,21	1,0979506	0,2884228
5	56	5963656,5	106493,87	0,7178016	0,9424105
6	54	5695104,4	105464,9	0,710866	0,9440988
7	52	5403992,2	103922,93	0,7004727	0,947928

Ao analisar, por Decomposição em Valor Singular aplicado a matriz de interação genótipo $\times$ ambiente que tem posto $p = \min(7, 58) = 7$, tem-se que $SQ_{G \times E}$ será decomposta em 7 componentes ortogonais.

Utilizando o método de Gollob (1968) para atribuir graus de liberdade, é realizado uma análise pelo teste F e os resultados são apresentados na Tabela 27. Ao nível de 1% de significância, nota-se que antes de corrigir os autovalores têm-se que o primeiro componente é significativo para o modelo e este corresponde a 24,90% da $SQ_{G \times E}$ utilizando 64 graus de liberdade (15,76%) do total de graus de liberdade da interação.

2.3.4.3 Análise AMMI com correção dos autovalores

Nota-se pelos resultados obtidos, na análise variância e na análise AMMI sem correção dos autovalores, que os resultados estão divergindo. Utilizando as sugestões de Muirhead (1987), obtêm-se os valores dos componentes corrigidos $\phi^{(1)}$, $\phi^{(2)}$ e $\phi^{(3)}$, que foram obtidos pelas expressões (26), (27) e (28), respectivamente, e são apresentados na Tabela 28.

Tabela 28 - Correção dos autovalores da matriz $(GE)(GE)^t$ do experimento 4 e os autovalores ajustados pela regressão isotônica

	λ^2	$\phi^{(1)}$	$\phi^{(1)*}$	$\phi^{(2)}$	$\phi^{(2)*}$	$\phi^{(3)}$	$\phi^{(3)*}$
1	16382511	12176016	12176016	12490712	12490712	12297661	12297661
2	12011970	7988810	10102268	8173070	10433605	9378255,4	9378255,4
3	10890890	11417058	10102268	11837723	10433605	8834243,1	8834243,1
4	9447805,9	12684960	10102268	13289821	10433605	7972661,3	7972661,3
5	5963656,5	3235033,8	7311882	3295635,3	7645047,1	5253732,1	5253732,1
6	5695104,4	8272709,6	7311882	8701160,7	7645047,1	5225869,6	5225869,6
7	5403992,2	-27249415	7311882	-23271583	7645047,1	5171576,6	5171576,6
Total	65795930		64418468		66726667		54133999
$\frac{\sum_{k=1}^7 \phi_k^{(1)*}}{\sum_{k=1}^{15} \lambda_k^2}$			0,9790		1,0141		0,8227

Pela Tabela 28, percebe-se que os novos valores das componentes não satisfazem algumas condições que deveriam ser respeitadas. A ordem decrescente dos valores não é verificada para nenhum dos métodos correção. Os valores de $\phi^{(1)}$ e $\phi^{(2)}$ apresentam outros problema assumindo valores negativos. Para superar esses problemas utiliza-se da regressão isotônica, mas propriamente utilizado o algoritmo de Steim apresentado por Lin e Perlman (1985). Os valores $\phi^{(1)*}$, $\phi^{(2)*}$ e $\phi^{(3)*}$, também são apresentados na Tabela 28 e se referem a regressão isotônica de $\phi^{(1)}$, $\phi^{(2)}$ e $\phi^{(3)}$, respectivamente.

Verifica-se ainda que $\sum_{k=1}^7 \phi_k^{(1)*}$ é equivale a 97,90% da $SQ_{G \times E}$ e o restante da $SQ_{G \times E}$ (2,10%) representa ruído detectado pela correção que estava presente na interação. O $\sum_{k=1}^7 \phi_k^{(2)*}$ é equivalente a 101,41% e 1,41% representa um acréscimo na $SQ_{G \times E}$ e $\sum_{k=1}^7 \phi_k^{(3)*}$ representa 82,27% da $SQ_{G \times E}$, sendo que a correção considerou 17,73% da interação de genótipos e ambientes como sendo ruído.

Pelas Tabelas 29, 30 e 31 que apresenta a análise de cada componente pelo teste F, com os graus de liberdade ajustados pelo método de Gollob (1968), para os autovalores corrigidos pelos métodos de correção 1, 2 e 3 (eq. (26), (27) e (28)). Considerando um nível de significância de 1%, tem-se que nenhum componente é significativo para o modelo, sendo assim esta correção considera que toda interação genótipo $\times$ ambiente é composta por ruídos, logo têm-se um modelo aditivo para análise dos dados.

Tabela 29 - Análise do Teste F para os componentes do Experimento 4 corrigidos pelo método 1

	GL	$SQ (\phi_k^{(1)*})$	QM	F	valor p
1	64	12176016	190250,26	1,2823456	0,0690009
2	62	10102268	162939,81	1,0982648	0,2830924
3	60	10102268	168371,14	1,1348736	0,2271039
4	58	10102268	174177,04	1,1740072	0,1774308
5	56	7311882	130569,32	0,8800776	0,7223915
6	54	7311882	135405,22	0,9126731	0,6547997
7	52	7311882	140613,12	0,9477759	0,5808778

Tabela 30 - Análise do Teste F para os componentes do Experimento 4 corrigidos pelo método 2

	GL	$SQ (\phi_k^{(2)*})$	QM	F	valor p
1	64	12490712	195167,37	1,3154885	0,0509088
2	62	10433605	168283,95	1,1342859	0,2250143
3	60	10433605	173893,41	1,1720954	0,1764162
4	58	10433605	179889,74	1,2125125	0,1346062
5	56	7645047,1	136518,7	0,9201783	0,6429453
6	54	7645047,1	141574,95	0,954259	0,5699088
7	52	7645047,1	147020,14	0,9909612	0,4933959

Tabela 31 - Análise do Teste F para os componentes do Experimento 4 corrigidos pelo método 1

	GL	SQ ($\phi_k^{(3)*}$)	QM	F	valor p
1	64	12297661	192150,95	1,2951569	0,0614554
2	62	9378255,4	151262,18	1,0195539	0,4357321
3	60	8834243,1	147237,38	0,9924255	0,4934448
4	58	7972661,3	137459,68	0,9265208	0,6328589
5	56	5253732,1	93816,644	0,6323532	0,9845605
6	54	5225869,6	96775,363	0,6522959	0,9761526
7	52	5171576,6	99453,397	0,6703467	0,9657076

Têm-se então que as conclusões obtidas com a análise AMMI após a correção dos autovalores, concordam com a análise da variância, ou seja, que o modelo aditivo é o melhor para analisar o conjunto de dados do experimento 4.

2.3.4.4 Número de Repetições

Tabela 32 - $RMSPD_{PRESS}$ e R_{AMMI} para o melhor modelo AMMI selecionado para os dados do Experimento 4 após a correção dos autovalores pelos métodos 1, 2 e 3

	$RMSPD_{PRESS}$	R_{AMMI}
Método 1	500,66856	1,4501439
Método 2	500,66856	1,4501439
Método 3	500,66856	1,4501439

São apresentados na Tabela 32 os valores das medidas $RMSPD_{PRESS}$ e R_{AMMI} , como para todos métodos de correção verificou-se que o modelo aditivo é o melhor então as medidas $RMSPD_{PRESS}$ e R_{AMMI} são todas iguais. Utilizando apenas o modelo aditivo têm-se um ganho de aproximadamente 2 repetições para o três métodos de correções e $RMSPD_{PRESS}$ de 500,66856.

2.3.5 Considerações gerais

Em todos os conjuntos de dados, após a correção dos autovalores, verifica-se que não existe a ordem decrescente de autovalores (para os métodos de correção 1, 2 e 3) e que algumas correções se tornam negativas (métodos de correção 1 e 2). Com isso é necessário que estas correções sejam ajustadas para que não ocorram valores negativos dos autovalores e nem tão pouco fora de ordenação.

Assim utilizou-se da técnica de regressão isotônica através do algoritmo de Stein.

Esta técnica se mostrou muito eficaz para realizar os ajuste necessários nos autovalores corrigidos em todos os conjuntos de dados.

Através do teste F , com os graus de liberdade ajustados pelo método de Gollob (1968), verificou-se a quantidade de componentes multiplicativos não nulos a serem retidos no modelo. O teste estatístico foi aplicado ao nível de 1% em vez de 5%, pois assim reduz-se a probabilidade de ocorrer o erro tipo I , isto é, de ajustar um modelo AMMI com um maior número de componentes, quando na verdade o modelo correto é mais parcimonioso. Contudo aumenta-se o probabilidade de ocorrer o erro tipo II , ou seja, de ajustar um modelo AMMI com menor número de componentes, sendo que o modelo correto é mais complexo.

A importância de controlar o erro tipo I é destacado por Cornelius, Seyedsar e Crossa (1992), que também consideram como sendo bastante liberal o método de Gollob (1968), podendo resultar em falsos resultados de significância e provavelmente uma seleção de maior número de componentes, aumentando assim a ocorrência do erro tipo I . Mas quando se aplicou o método de Gollob (1968) com os métodos de correção para autovalores, verificou-se que houve uma redução no número de componentes retidas no modelo em todos conjuntos de dados, com exceção no experimento 2, que manteve o mesmo número de componente.

Os métodos de correções mostraram-se bastantes distintos quando se compara as taxas de correções. O método 3 mostrou-se bastante rigoroso e apresentou as maiores taxas de correções, ou seja, antes de aplicar o teste F para as componentes, a correção descartou grande parte da $SQ_{G \times E}$, considerando essa parte que foi descartado como ruído. O mesmo método considerou no máximo 41,62% da soma de quadrados da interação genótipos $\times$ ambiente (no experimento 3), quando esta interação foi considerada significativa pela ANAVA e quando a interação genótipo $\times$ ambiente foi considerada não significativa (experimento 4), o método 3 considerou 82,27%. O método 2 foi o que se mostrou menos rigoroso e apresentou as menores taxas de correções, sendo que no experimento 2 este método considerou 79,18% (maior taxa de correção do método 2) da $SQ_{G \times E}$. Já o método 1 mostrou-se intermediário aos métodos 2 e 3, mas apresentou resultados próximo do método 2.

A medida $RMSPD_{PRESS}$ reflete os desvios das predições do modelo AMMI para cada combinação de genótipos e ambientes, assim a preferência deve recair sobre o modelo cuja diferença preditiva seja menor. Pelos valores da $RMSPD_{PRESS}$ obtida para cada método de correção e em cada experimento, verifica-se que em todos os experimentos a menor medida $RMSPD_{PRESS}$ é

obtida do modelo ajustado para os autovalores corrigidos pelo método 2, mas ficou muito próximo das medidas obtidas quando se utilizou o método 1 para corrigir os autovalores. Quando se utiliza o método 3 para corrigir os autovalores e ajusta um modelo AMMI, verifica-se os maiores valores para a medida $RMSPD_{PRESS}$. Nota-se então que existe uma relação entre o valor da medida $RMSPD_{PRESS}$ e a rigorosidade dos métodos de correção para autovalores, ou seja, quanto mais rigoroso for o método de correção, maior será a medida $RMSPD_{PRESS}$.

Através da medida R_{AMMI} , estimou-se o ganho em termos do número de repetições ao se fazer a correção dos autovalores dos modelos AMMI. Verificou-se através da medida R_{AMMI} que o método 2 se mostrou melhor, fornecendo um benefício de repetições sempre próximo de três. Já o método de correção 3 é o que apresenta o menor ganho em número de repetições, na maioria dos experimentos esteve próximo de duas repetições. Verifica-se também que existe uma tendência de relação inversa entre a medida R_{AMMI} e $RMSPD_{PRESS}$, ou seja, quanto menor for o valor da $RMSPD_{PRESS}$ maior foi o ganho em número de repetições. Logo esta mesma interpretação vale para a rigorosidade do método de correção, quanto menos rigoroso for o método de correção maior foi o ganho em termos de número de repetições.

3 Conclusões

Os resultados obtidos nessa pesquisa permitem extrair as seguintes conclusões:

- a) A regressão isotônica juntamente com o algoritmo de Steim foi necessária em todos conjuntos de dados. Esta ferramenta se mostrou muito importante para transformar as correções negativas em positivas e para a ordenação dos autovalores corrigidos;
- b) Em relação ao número de componentes retidas em cada modelo, percebe-se que cada conjunto dos dados, uma redução no número de componentes significativos para serem retidos nos modelos, fazendo com que os modelos AMMI selecionados sejam mais parcimoniosos quando se utiliza qualquer um dos métodos de correção
- c) O método 2 apresenta a menor taxa de correção, considerando no mínimo 79,18% do total da soma de quadrados da interação genótipo $\times$ ambiente ;
- d) O método 3 apresenta a maior taxa de correção, considerando no máximo 41,62% do total da soma de quadrados da interação genótipo $\times$ ambiente, quando a interação foi significativa, mas chegou a 82,27% do total da soma de quadrados da interação genótipo $\times$ ambiente, quando a interação foi não significativa;
- e) Em relação a medida $RMSPD_{PRESS}$, os menores valores foram obtidos quando se utilizou o método de correção 2, mas estes ficaram muito próximos das medidas obtidas quando se utilizou o método de correção 1. Já o método de correção 3 apresentou os maiores valores para $RMSPD_{PRESS}$;
- f) O método 2 também se mostrou melhor quando o interesse era verificar o ganho em número de repetições, sendo que este benefício esteve sempre próximo de 3 repetições. Já o método 3 é o que apresenta um menor ganho em número de repetições, na maioria dos experimentos teve apenas duas repetição de benefício.

REFERÊNCIAS

- ALLARD, R.W. **Princípios do melhoramento genético das plantas**. Rio de Janeiro: USAID/Edgard Blucher, 1971. 381p.
- ALLARD, R.W. BRADSHAW, A.D. Implications of genotype-environmental interactions in applied plant breeding. **Crop Science**, Madison, v.4, n.5, p.503-508, 1964.
- ALLEN, D.M. Mean squares error of prediction as a criterion for selecting variables. **Technometrics**, Alexandria, v.13, p.469-475, 1971.
- ANDERSON, T.W. **An introduction to multivariate statistical analysis**. New York: Wiley, 1958. 374p.
- ANNICCHIARICO, P. Additive main effects and multiplicative interaction (AMMI) analysis of genotype-location interactions in variety trials repeated over years. **Theoretical and Applied Genetics**, New York, v.94, p.1072-1077, 1997a.
- ANNICCHIARICO, P. Joint regression vs AMMI analysis of genotype-environment interactions for cereals in Italy. **Euphytica**, Wageningen, v.94, p.53-62, 1997b.
- ARAÚJO, L.B.; DIAS, C.T.S. Viés e distribuição de autovalores de matrizes de variâncias e covariâncias nos modelos AMMI. In: ESCOLA DE MODELOS DE REGRESSÃO, 9., 2005, São Pedro. **Anais...**, São Pedro: ABE, 2005, p.60.
- AYER, M.; BRUNK, H.; EWING, G.; REID, W.; SILVERMAN, E. An empirical distribution function for sampling with incomplete information. **Annals of Mathematical Statistics**, Ann Arbor, v.26, n.4, p.641-647, 1955.
- BOLDRINI, J.L.; COSTA, S.I.R.; RIBEIRO, V.L.F.; WETZLER, H.G. **Algebra Linear**. São Paulo: Harper e Row do Brasil, 1980. 411p.
- CHAVES, J.L. Interação de cultivares com ambientes. In: NASS, L.L.; VALOIS, A.C.C.; MELO, I.S.; VALADARES, M.C. **Recursos genéticos e melhoramento - plantas**. Rondonópolis: Fundação MT, 2001, p.673-713
- CORNELIUS, P.L.; CROSSA, J. Prediction assessment of shrinkage estimators of multiplicative models for multi-environment trials. **Crop Science**, Madison, v.39, p.998-1009, 1999.

CORNELIUS, P.L.; CROSSA, J.; SEYEDSADR, M.S. Teste and estimators of multiplicative models for variety trials. In: ANNUAL KANSAS STATE UNIVERSITY CONFERENCE ON APPLIED STATISTICS IN AGRICULTURE, 5., 1993, Manhattan. **Annals...** Manhattan: Milliken, 1993, p.156-166.

CORNELIUS, P.L.; SEYEDSADR, M.; CROSSA, J. Using the shifted multiplicative model to search for “ separability ” in crop cultivar trials. **Theoretical Applied Genetics**, New York, v.84, p.161-172, 1992.

CROSSA, J. Statistical Analyses of multilocation trials. **Advances in Agronomy**, New York, v.44, p.55-85, 1990.

CRUZ, C.D.; REGAZZI, A.J. **Modelos biométricos aplicados ao melhoramento genético**. Viçosa: UFV, 2001. 390p.

CRUZ, C.D.; CARNEIRO, P.C.S. **Modelos biométricos aplicados ao melhoramento genético**. Viçosa: UFV, 2003. 585p.

DIAS, C.T.S.; KRZANOWSKI, W.J. Model selection and cross validation in additive main effect and multiplicative interaction models. **Crop Science**, Madison, v.43, p.865-873, 2003.

DUARTE, J.B.; VENCOSKY, R. **Interação genótipo x ambiente: uma introdução à análise “AMMI”**. Ribeirão Preto: Sociedade Brasileira de Genética, 1999. 60p. (Série Monografias, 9).

EBERHART, S.A.; RUSSELL, W.A. Stability parameters for comparing varieties. **Crop Science**, Madison, v.6, p.36-40, 1966.

FALCONER, D.S. **Introdução à genética quantitativa**. Trad. de Almeida, M. e Silva, J.C. Viçosa: UFV, 1987. 279p.

FEHR, W.R. **Principles of cultivar development**. New York: Macmillan, 1987, cap.18, p.247-258.

FINLAY, K.W.; WILKINSON, G.N. The analysis of adaptation in plant-breeding programe. **Australian Journal of Agricultural research**, Melbourne, v.14, n.6, p.117-127, 1963.

FOX, P.N.; CROSSA, J.; ROMAGOSA, I. Multi-environment testing and genotype $\times$ environment interaction. In: KEMPTON, R.A.; FOX, P.N. **Statistical methods for plant variety evaluation**. New York: Chapman e hall, 1997, cap.8, p.117-138.

GABRIEL, K.R. The biplot graphic display of matrices with application to principal component analysis. **Biometrika**, Cambridge, v.58, p.453-467, 1971.

GAUCH, H.G. Model selections and validation for yield trials with interactions. **Biometrics**, Washington, v.44, p.705-715, 1988.

GAUCH JÚNIOR, H.G. **Statistical analysis of regional yield trials: AMMI analysis of factorial designs**. Amsterdam: Elsevier, 1992, 277p.

GAUCH JÚNIOR, H.G.; ZOBEL, R.W. Predictive and postdictive success of statistical analysis of yield trials. **Theoretical and Applied Genetics**, New York, v.76, p.1-10, 1988.

GAUCH JÚNIOR, H.G.; ZOBEL, R.W. AMMI analysis of yield trials. In: KANG, M.S.; GAUCH JÚNIOR, H.G. **Genotype-by-environment interaction** Boca Raton: CRC Press, 1996. cap.4. p.85-122.

GOLLOB, H.F. A statistical model which combines features of factor analytic and analysis of variance techniques. **Psychometrika**, New York, v.33, p.73-115, 1968.

GONZÁLEZ, L.M.A.R. **Estudo da interação de genótipos $\times$ ambientes para a cultura do algodão (*Gossypium* sp.) na região oriental do Paraguai**. 1988. 91p. Dissertação (Mestrado em Genética e Melhoramento de Plantas) - Escola Superior de Agricultura "Luiz de Queiroz", Universidade de São Paulo, Piracicaba, 1988.

HILL, R.R.; ROSENBERGER, J.L. Methods for combining data from germplasm evaluation trials. **Crop Science**, Madison, v.25, p.467-470, 1985.

HOOGERHEIDE, E.S.S. **Estabilidade fenotípica de cultivares herbáceo em diferentes sistemas de produção no Estado do Mato Grosso**. 2004. 80p. Dissertação (Mestrado em Genética e Melhoramento de Plantas) - Escola Superior de Agricultura "Luiz de Queiroz", Universidade de São Paulo, Piracicaba, 2004.

JOHNSON, R.A.; WICHERN, D.W. **Applied multivariate statistical analysis**. Madison: Prentice Hall International, 1998. 816p.

JOHNSTONE, I.M. On the distribution of the largest eigenvalue in principal components analysis. **The Annals of Statistics**, Hayward, v.29, n.2, p.295-327, 2001.

KANG, M.S. Using genotype-by-environment interaction for crop cultivar development. **Advances in Agronomy**, New York, v.62, p.199-252, 1998.

KANG, M.S; MAGARI, R. New developments in selecting for phenotypic stability in crop breeding. In: KANG, M.S.; GAUCH, H.G. **Genotype-by-environment interaction**. Boca Raton: CRC Press, 1996, cap.1, p.1-14.

KEMPTON, R.A. The use of biplots in interpreting variety by environmet interactions. **Journal of Agricultural Science**, Cambridge, v.103, p.123-135, 1984.

LANCZOS, C. **Linear differential operators**. New York : Van Nostrand, 1961. 564p.

LIN, S.P.; PERLMAN, M.D. A Monte Carlo comparison of four estimators of a covariance matrix. In: KRISHNAIAH, P.R.**Multivariate Analysis - VI**. Amsterdam: Elsevier, 1985, cap.26, p.411-429.

MANDEL, J. The partitioning of interactions in analysis of variance. **Journal of Research of the National Bureau of Standards , Series B**, Washington, v.73, p.309-328, 1969.

MANDEL, J. A new analysis of variance model for non-adittive data. **Technometrics**, Alexandria, v.13, n.1, p.1-18, 1971.

MANDEL, J. Use of the singular value decomposition in regression analysis. **The American Statistician**, Washington, v.36, n.1, p.15-24, 1982.

MIRANDA, F.T.S. **Métodos de regressão e uni-multivariados para a redução do numero de repetições em experimentos intermediários de um programa de melhoramento de soja**. 2004. 110 p. Tese (Doutorado em Genética e Melhoramento de Plantas) - Escola Superior de Agricultura “Luiz de Queiroz”, Universidade de São Paulo. Piracicaba, 2004.

MORENO-GONZÁLEZ, J.; CROSSA, J.; CORNELIUS, P.L. Additive main effects and multiplicative interaction model: I. Theory on variance components for predicting cell means. **Crop Science**, Madison, v.43, p.1967-1975, 2003a.

MORENO-GONZÁLEZ, J.; CROSSA, J.; CORNELIUS, P.L. Additive main effects and multiplicative interaction model: II. Theory on shrinkage factors for predicting cell means. **Crop Science**, Madison, v.43, p.1976-1982, 2003b.

MUIRHEAD, R.J. Developments in eigenvalue estimation. In: GUPTA, A. K. **Advances in Multivariate Statistical**. Boston: Reidel, 1987, p.277-288.

NACHIT, M.M.; NACHIT, G.; KETATA, H.; GAUCH, H.G.; ZOBEL, R.W. Use of AMMI and linear regression models to analyze genotype-environment interaction in durum wheat. **Theoretical and Applied genetics**, New York, v.83, p.597-601, 1992.

OLIVEIRA, A.B.; DUARTE, J.B.; PINHEIRO, J.B. Emprego da análise AMMI na avaliação da estabilidade produtiva em soja. **Pesquisa Agropecuária Brasileira**, Brasília, v.38, n.3, p.357-364, 2003.

PIEPHO, H.-P. Best Linear Unbiased Prediction (BLUP) for regional yield trials: a comparison to additive main effects and multiplicative interactions (AMMI) analysis. **Theoretical and Applied genetics**, New York, v.89, p.647-654, 1994.

PIKE, E.W.; SILVERBERG, T.R. Designing mechanical computers. **Machine Design**, Cleveland, v.24, p.131-137, 159-163, 1952.

PIMENTEL-GOMES, F. **Curso de estatística experimental**. Piracicaba, 14ed. 2000. 477p.

RAMALHO, M.A.P.; SANTOS, J.B.; ZIMMERMANN, M.J.O. **Genética quantitativa em plantas autógamas: aplicações ao melhoramento do feijoeiro**. Goiânia: UFG, 1993. 271p.

ROBERTSON, T.; WRIGHT, F. T.; DYKSTRA, R. L. **Order Restricted Statistical Inference**. New York: Wiley, 1988, 521p.

SAS Institute inc. **SAS 9.2**. Cary, 2003.

SILVA, J.G.C.; BARRETO, J.N. Aplicação de regressão linear segmentada em estudos da interação genótipo $\times$ ambiente. In: SIMPÓSIO DE ESTATÍSTICA APLICADA À EXPERIMENTAÇÃO AGRONÔMICA, 1., 1985, Campinas. **Resumo...** Campinas: Fundacao Cargill, 1985, p.49-50.

STROUP, W.W.; MULITZE, D.K. Nearest neighbor adjusted best linear unbiased prediction. **The American Statistician**, Washington, v.45, p.194-200, 1991.

VENCOVSKY, R.; BARRIGA, P. **Genética biométrica no fitomelhoramento**. Ribeirão Preto: Sbg, 1992. 486p.

WEBER, W.E.; WRICKE, G.; WESTERMANN, T. Selection of genotypes and prediction of performance by analysis genotype-by-environment interactions. In: KANG, M.S; GAUCH, H.G. **Genotype by environment interactions**. Boca Raton: CRC Press, 1996, cap.13, p.353-371.

WILLIAMS, E.J. The interpretation of interactions in factorial experiments. **Biometrika**, Cambridge, v.39, p.65-81, 1952.

ZOBEL, R.W.; WRIGHT M.J.; GAUCH JÚNIOR., H.G Statistical analysis of a yield trial. **Agro-nomy Journal**, Madison, v.80 , p.388-393, 1988.

ANEXOS

```
#####
* Análise da variância dos experimento;
#####
proc glm data=TRIAL1 outstat=anavaexp;
class REP SITE GEN;
model YIELD=REP SITE GEN SITE*GEN/ss1;
run;
data anavaexp2;
set anavaexp;
if _source_ = 'ERROR';
SQRes=ss;
gl_Res=df;
QMRes=SQRes/gl_Res;
keep gl_Res QMRes;
data anavaexp3;
set anavaexp;
if _source_ = 'REP';
SQBloco=ss;
gl_bloco=df;
QMBloco=SQBloco/gl_bloco;
nRep=df+1;
keep gl_bloco
QMBloco nrep;
#####
*Obter a tabela de dupla entrada residual (GE);
#####
proc sort data=TRIAL1;
by site gen;
run;
proc means data=TRIAL1 noprint;
by site gen;
var YIELD;
output out = medias mean=YIELDM;
run;
proc sort data=medias;
by gen site;
run;
proc transpose data=medias out = medias2;
by gen;
id site;
var YIELDM;
run;
#####
* Uso do módulo de matrizes;
#####
proc iml;
use anavaexp2;
read var {QMRes} into QMRes;
read var {gl_Res} into gl_Res;
```

```

use anavaexp3;
read var {gl_bloco} into gl_bloco;
read var {QMBloco} into QMBloco;
read var {nRep} into nRep;
use medias2;
read all into YIELDM;
Y=YIELDM[,2:ncol(YIELDM)];
g=nrow(Y);
e=ncol(Y);
Yi=j(g,e,0); do i=1 to g;
    Yi[i,]=j(1,e,sum(Y[i,])/e);
end;
Yj=j(g,e,0); do j=1 to e;
    Yj[,j]=j(g,1,sum(Y[,j])/g);
end;
Y_barra = j(g,e,sum(Y)/(g*e));
GE = Y - Yi - Yj + Y_barra;
SQGen=e*sum((Yi[,1] - Y_barra[,1])##2); SQSite=g*sum((Yj[1,] -
Y_barra[1,])##2); SQInt=sum(GE##2);
p=min(g-1,e-1);
call SVD (U,L,V,GE);
L_auto=L[1:p,];
#####
* Análise AMMI método de Gollob;
#####
SQ_gollob=j(p,1,0);
gl_gollob=j(p,1,0);
QM_gollob=j(p,1,0);
F_gollob=j(p,1,0);
p_valor_gollob=j(p,1,0);
do i=1 to p;
    SQ_gollob[i,]=(L_auto[i,]##2);
    gl_gollob[i,]=g + e - 1 - 2*i;
    QM_gollob[i,]=SQ_gollob[i,]/gl_gollob[i,];
    F_gollob[i,]=QM_gollob[i,]/(QMRes/nRep);
    p_valor_gollob[i,]=1-probf(F_gollob[i,],gl_gollob[i,],gl_Res);
    if p_valor_gollob[i,] < 0.0001 then
        do;
            p_valor_gollob[i,]= 0;
        end;
    else
        p_valor_gollob[i,]=p_valor_gollob[i,];
    end;
end;
print "Análise AMMI pelo método de Gollob";
print gl_gollob SQ_gollob QM_gollob F_gollob p_valor_gollob;
#####
* Correção dos aultovalores;
#####
lambda2=L_auto##2;

```

```
#####
*Método de correção 1;
#####
FI_1=j(p,1,0);
nlambda_1=e*lambda2;
alpha_1=j(p,1,0);
#####
*Método de correção 2;
#####
FI_2=j(p,1,0);
nlambda_2=e*lambda2;
alpha_2=j(p,1,0);
#####
*para i = 1;
#####
li=j(p-1,1,lambda2[1,]);
lj=lambda2[2:p,];
alpha_1[1,]=(e-g+1+2*lambda2[1,]*sum((li-lj)##-1));
FI_1[1,]=nlambda_1[1,]/alpha_1[1,];
alpha_2[1,]=(e-g-1+2*lambda2[1,]*sum((li-lj)##-1));
FI_2[1,]=nlambda_2[1,]/alpha_2[1,];
#####
*para i=2 ate p-1;
#####
do i=2 to (p-1);
    li=j(p-1,1,lambda2[i,]);
    lj=lambda2[1:i-1,]/lambda2[i+1:p,];
    alpha_1[i,]=(e-g+1+2*lambda2[i,]*sum((li-lj)##-1));
    FI_1[i,]=nlambda_1[i,]/alpha_1[i,];
    alpha_2[i,]=(e-g-1+2*lambda2[i,]*sum((li-lj)##-1));
    FI_2[i,]=nlambda_2[i,]/alpha_2[i,];
end;
#####
*para i = p;
#####
li=j(p-1,1,lambda2[p,]);
lj=lambda2[1:(p-1),];
alpha_1[p,]=(e-g+1+2*lambda2[p,]*sum((li-lj)##-1));
FI_1[p,]=nlambda_1[p,]/alpha_1[p,];
alpha_2[p,]=(e-g-1+2*lambda2[p,]*sum((li-lj)##-1));
FI_2[p,]=nlambda_2[p,]/alpha_2[p,];
#####
*Método de correção 3;
#####
FI_3=j(p,1,0);
alpha_3=j(p,1,0);
nlambda_3=j(p,1,0);
b=6*(g-2)/(5*((e+g-1)##2));
c=5*8*((g-2)##2)/((e+g-1)##2);
```

```

ln_li=log(e*lambda2[1:p,]);
do i=1 to p;
  alpha_3[i,]=e*lambda2[i,]*( (b + sum(ln_li##2)) -
  c*ln_li[i,]*(e+g+1-2*i) );
  nlambda_3[i,]=(e+g+1-2*i)*(b+sum(ln_li##2));
  FI_3[i,]=alpha_3[i,]/nlambda_3[i,];
end;
print lambda2 FI_1 FI_2 FI_3;
#####
* Regressão Isotonica dos autovalores corrigidos pelo método 1
#####
#####
*Passo 1;
#####
do i=1 to p;
  if alpha_1[p+1-i,] < 0 then
    do;
      alpha_1[p+1-i,]=(alpha_1[p+1-i,]+alpha_1[p-i,]);
      nlambda_1[p+1-i,]=(nlambda_1[p+1-i,]+nlambda_1[p-i,]);
      alpha_1[p-i,]=alpha_1[p+1-i,];
      nlambda_1[p-i,]=nlambda_1[p+1-i,];
    end;
  else
    alpha_1[p+1-i,]=alpha_1[p+1-i,];
    nlambda_1[p+1-i,]=nlambda_1[p+1-i,];
  end;
do i=1 to p;
  if alpha_1[i,] > 0 then
    do;
      troca_alpha_1=alpha_1[i,];
      troca_lambda_1=nlambda_1[i,];
    end;
  else
    alpha_1[i,]=troca_alpha_1;
    nlambda_1[i,]=troca_lambda_1;
  end;
end;
Fi_1_iso=j(p,1,0);
do i=1 to p;
  Fi_1_iso[i,]=nlambda_1[i,]/alpha_1[i,];
end;
#####
*Passo 2;
#####
do i=1 to p-1;
  if Fi_1_iso[p-i+1,] > Fi_1_iso[p-i,] then
    do;
      alpha_1[p-i,]=(alpha_1[p-i+1,]+alpha_1[p-i,]);
      nlambda_1[p-i,]=(nlambda_1[p-i+1,]+nlambda_1[p-i,]);
      alpha_1[p-i+1,]=alpha_1[p-i,];
    end;
  end;
end;

```

```

        nlambda_1[p-i+1,]=nlambda_1[p-i,];
        Fi_1_iso[p-i,]=nlambda_1[p-i,]/alpha_1[p-i,];
        Fi_1_iso[p-i+1,]=nlambda_1[p-i+1,]/alpha_1[p-i+1,];
    end;
else
    alpha_1[p-i+1,]=alpha_1[p-i+1,];
    alpha_1[p-i,]=alpha_1[p-i,];
    nlambda_1[p-i+1,]=nlambda_1[p-i+1,];
    nlambda_1[p-i,]=nlambda_1[p-i,];
    Fi_1_iso[p-i,]=nlambda_1[p-i,]/alpha_1[p-i,];
    Fi_1_iso[p-i+1,]=nlambda_1[p-i+1,]/alpha_1[p-i+1,];
end;
do i=1 to (p-1);
    if Fi_1_iso[i,] >= Fi_1_iso[i+1,] then
        do;
            Fi_1_iso[i,]=Fi_1_iso[i,];
            Fi_1_iso[i+1,]=Fi_1_iso[i+1,];
            nlambda_1[i,]=nlambda_1[i,];
            nlambda_1[i+1,]=nlambda_1[i+1,];
            alpha_1[i,]=alpha_1[i,];
            alpha_1[i+1,]=alpha_1[i+1,];
        end;
    else;
        Fi_1_iso[i+1,]=Fi_1_iso[i,];
        nlambda_1[i+1,]=nlambda_1[i,];
        alpha_1[i+1,]=alpha_1[i,];
    end;
end;
#####
* Regressão Isotonica dos autovalores corrigidos pelo método 2
#####
#####
*Passo 1;
#####
do i=1 to p;
    if alpha_2[p+1-i,] < 0 then
        do;
            alpha_2[p+1-i,]=(alpha_2[p+1-i,]+alpha_2[p-i,]);
            nlambda_2[p+1-i,]=(nlambda_2[p+1-i,]+nlambda_2[p-i,]);
            alpha_2[p-i,]=alpha_2[p+1-i,];
            nlambda_2[p-i,]=nlambda_2[p+1-i,];
        end;
    else
        alpha_2[p+1-i,]=alpha_2[p+1-i,];
        nlambda_2[p+1-i,]=nlambda_2[p+1-i,];
    end;
end;
do i=1 to p;
    if alpha_2[i,] > 0 then
        do;
            troca_alpha_2=alpha_2[i,];

```

```

        troca_lambda_2=nlambda_2[i,];
    end;
else
    alpha_2[i,]=troca_alpha_2;
    nlambda_2[i,]=troca_lambda_2;
end;
Fi_2_iso=j(p,1,0);
do i=1 to p;
    Fi_2_iso[i,]=nlambda_2[i,]/alpha_2[i,];
end;
#####
*Passo 2;
#####
do i=1 to p-1;
    if Fi_2_iso[p-i+1,] > Fi_2_iso[p-i,] then
        do;
            alpha_2[p-i,]=(alpha_2[p-i+1,]+alpha_2[p-i,]);
            nlambda_2[p-i,]=(nlambda_2[p-i+1,]+nlambda_2[p-i,]);
            alpha_2[p-i+1,]=alpha_2[p-i,];
            nlambda_2[p-i+1,]=nlambda_2[p-i,];
            Fi_2_iso[p-i,]=nlambda_2[p-i,]/alpha_2[p-i,];
            Fi_2_iso[p-i+1,]=nlambda_2[p-i+1,]/alpha_2[p-i+1,];
        end;
    else
        alpha_2[p-i+1,]=alpha_2[p-i+1,];
        alpha_2[p-i,]=alpha_2[p-i,];
        nlambda_2[p-i+1,]=nlambda_2[p-i+1,];
        nlambda_2[p-i,]=nlambda_2[p-i,];
        Fi_2_iso[p-i,]=nlambda_2[p-i,]/alpha_2[p-i,];
        Fi_2_iso[p-i+1,]=nlambda_2[p-i+1,]/alpha_2[p-i+1,];
    end;
do i=1 to (p-1);
    if Fi_2_iso[i,] >= Fi_2_iso[i+1,] then
        do;
            Fi_2_iso[i,]=Fi_2_iso[i,];
            Fi_2_iso[i+1,]=Fi_2_iso[i+1,];
            nlambda_2[i,]=nlambda_2[i,];
            nlambda_2[i+1,]=nlambda_2[i+1,];
            alpha_2[i,]=alpha_2[i,];
            alpha_2[i+1,]=alpha_2[i+1,];
        end;
    else;
        Fi_2_iso[i+1,]=Fi_2_iso[i,];
        nlambda_2[i+1,]=nlambda_2[i,];
        alpha_2[i+1,]=alpha_2[i,];
    end;
end;
#####
* Regressão Isotonica dos autovalores corrigidos pelo método 3;
#####

```



```

#####
*Passo 1;
#####
do i=1 to p;
  if alpha_3[p+1-i,] < 0 then
    do;
      alpha_3[p+1-i,]=(alpha_3[p+1-i,]+alpha_3[p-i,]);
      nlambda_3[p+1-i]=(nlambda_3[p+1-i,]+nlambda_3[p-i,]);
      alpha_3[p-i,]=alpha_3[p+1-i,];
      nlambda_3[p-i]=nlambda_3[p+1-i,];
    end;
  else
    alpha_3[p+1-i,]=alpha_3[p+1-i,];
    nlambda_3[p+1-i,]=nlambda_3[p+1-i,];
  end;
do i=1 to p;
  if alpha_3[i,] > 0 then
    do;
      troca_alpha_3=alpha_3[i,];
      troca_lambda_3=nlambda_3[i,];
    end;
  else
    alpha_3[i,]=troca_alpha_3;
    nlambda_3[i,]=troca_lambda_3;
  end;
Fi_3_iso=j(p,1,0);
do i=1 to p;
  Fi_3_iso[i,]=alpha_3[i,]/nlambda_3[i,];
end;
#####
*Passo 2;
#####
do i=1 to p-1;
  if Fi_3_iso[p-i+1,] > Fi_3_iso[p-i,] then
    do;
      alpha_3[p-i,]=(alpha_3[p-i+1,]+alpha_3[p-i,]);
      nlambda_3[p-i,]=(nlambda_3[p-i+1,]+nlambda_3[p-i,]);
      alpha_3[p-i+1,]=alpha_3[p-i,];
      nlambda_3[p-i+1,]=nlambda_3[p-i,];
      Fi_3_iso[p-i,]=alpha_3[p-i,]/nlambda_3[p-i,];
      Fi_3_iso[p-i+1,]=alpha_3[p-i+1,]/nlambda_3[p-i+1,];
    end;
  else
    alpha_3[p-i+1,]=alpha_3[p-i+1,];
    alpha_3[p-i,]=alpha_3[p-i,];
    nlambda_3[p-i+1,]=nlambda_3[p-i+1,];
    nlambda_3[p-i,]=nlambda_3[p-i,];
    Fi_3_iso[p-i,]=alpha_3[p-i,]/nlambda_3[p-i,];
    Fi_3_iso[p-i+1,]=alpha_3[p-i+1,]/nlambda_3[p-i+1,];
  end;

```

```

end;
do i=1 to (p-1);
  if Fi_3_iso[i,] >= Fi_3_iso[i+1,] then
    do;
      Fi_3_iso[i,]=Fi_3_iso[i,];
      Fi_3_iso[i+1,]=Fi_3_iso[i+1,];
      nlambda_3[i,]=nlambda_3[i,];
      nlambda_3[i+1,]=nlambda_3[i+1,];
      alpha_3[i,]=alpha_3[i,];
      alpha_3[i+1,]=alpha_3[i+1,];
    end;
  else;
    Fi_3_iso[i+1,]=Fi_3_iso[i,];
    nlambda_3[i+1,]=nlambda_3[i,];
    alpha_3[i+1,]=alpha_3[i,];
  end;
end;
print lambda2 FI_1 FI_1_iso FI_2 FI_2_iso FI_3 FI_3_iso;
#####
*Análise AMMI para os autovalores corrigidos;
#####
*Analise AMMI para os autovalores corrigidos pelo metodo 1;
#####
SQ_gollob_1=j(p,1,0);
gl_gollob_1=j(p,1,0);
QM_gollob_1=j(p,1,0);
F_gollob_1=j(p,1,0);
p_valor_gollob_1=j(p,1,0);
do i=1 to p;
  SQ_gollob_1[i,]=FI_1_iso[i,];
  gl_gollob_1[i,]=g + e - 1 - 2*i;
  QM_gollob_1[i,]=SQ_gollob_1[i,]/gl_gollob_1[i,];
  F_gollob_1[i,]=QM_gollob_1[i,]/(QMRes/nRep);
  p_valor_gollob_1[i,]=1-probf(F_gollob_1[i,],gl_gollob_1[i,],gl_Res);
  if p_valor_gollob_1[i,] < 0.0001 then
    do;
      p_valor_gollob_1[i,]= 0;
    end;
  else
    p_valor_gollob_1[i,]=p_valor_gollob_1[i,];
  end;
end;
print "Analise AMMI pelo método de Gollob, para os autovalores
corrigidos pelo metodo 1";
print gl_gollob_1 SQ_gollob_1 QM_gollob_1 F_gollob_1 p_valor_gollob_1;
#####
*Analise AMMI para os autovalores corrigidos pelo metodo 2;
#####
SQ_gollob_2=j(p,1,0);
gl_gollob_2=j(p,1,0);

```

```

QM_gollob_2=j(p,1,0);
F_gollob_2=j(p,1,0);
p_valor_gollob_2=j(p,1,0);
do i=1 to p;
    SQ_gollob_2[i,]=FI_2_iso[i,];
    gl_gollob_2[i,]=g + e - 1 - 2*i;
    QM_gollob_2[i,]=SQ_gollob_2[i,]/gl_gollob_2[i,];
    F_gollob_2[i,]=QM_gollob_2[i,]/(QMRes/nRep);
    p_valor_gollob_2[i,]=1-probf(F_gollob_2[i,],gl_gollob_2[i,],gl_Res);
    if p_valor_gollob_2[i,] < 0.0001 then
        do;
            p_valor_gollob_2[i,]= 0;
        end;
    end;
end;
print "Analise AMMI pelo método de Gollob,para os autovalores
corrigidos pelo metodo 2";
print gl_gollob_2 SQ_gollob_2 QM_gollob_2 F_gollob_2 p_valor_gollob_2;
#####
*Analise AMMI para os autovalores corrigidos pelo metodo 3;
#####
SQ_gollob_3=j(p,1,0);
gl_gollob_3=j(p,1,0);
QM_gollob_3=j(p,1,0);
F_gollob_3=j(p,1,0);
p_valor_gollob_3=j(p,1,0);
do i=1 to p;
    SQ_gollob_3[i,]=FI_3[i,];
    gl_gollob_3[i,]=g + e - 1 - 2*i;
    QM_gollob_3[i,]=SQ_gollob_3[i,]/gl_gollob_3[i,];
    F_gollob_3[i,]=QM_gollob_3[i,]/(QMRes/nRep);
    p_valor_gollob_3[i,]=1-probf(F_gollob_3[i,],gl_gollob_3[i,],gl_Res);
    if p_valor_gollob_3[i,] < 0.0001 then
        do;
            p_valor_gollob_3[i,]= 0;
        end;
    else
        p_valor_gollob_3[i,]=p_valor_gollob_3[i,];
    end;
end;
print "Analise AMMI pelo método de Gollob, para os autovalores
corrigidos pelo metodo 3";
print gl_gollob_3 SQ_gollob_3 QM_gollob_3 F_gollob_3 p_valor_gollob_3;
#####
*Predição do modelo e a medida PRESS;
#####
L_gollob_1=j(nrow(L),1,0);
do i=1 to p;
    if p_valor_gollob_1[i,]> 0.01 then
        do;
            L_gollob_1[i,]=0;
        end;
    end;
end;

```

```

        end;
    else
        L_gollob_1[i,]=SQ_gollob_1[i,]##(1/2);
    end;
Y_HAT_GOLLOB_1=Yi+Yj-Y_barra + U*diag(L_gollob_1)*t(V);
res_gollob_1=Y-Y_HAT_GOLLOB_1;
press_gollob_1=sum(res_gollob_1##2);
L_gollob_2=j(nrow(L),1,0);
do i=1 to p;
    if p_valor_gollob_2[i,]> 0.01 then
        do;
            L_gollob_2[i,]=0;
        end;
    else
        L_gollob_2[i,]=SQ_gollob_2[i,]##(1/2);
    end;
Y_HAT_GOLLOB_2=Yi+Yj-Y_barra + U*diag(L_gollob_2)*t(V);
res_gollob_2=Y-Y_HAT_GOLLOB_2;
press_gollob_2=sum(res_gollob_2##2);
L_gollob_3=j(nrow(L),1,0);
do i=1 to p;
    if p_valor_gollob_3[i,]> 0.01 then
        do;
            L_gollob_3[i,]=0;
        end;
    else
        L_gollob_3[i,]=SQ_gollob_3[i,]##(1/2);
    end;
Y_HAT_GOLLOB_3=Yi+Yj-Y_barra + U*diag(L_gollob_3)*t(V);
res_gollob_3=Y-Y_HAT_GOLLOB_3;
press_gollob_3=sum(res_gollob_3##2);
#####
*RMSPD_PRESS e R_ammi;
#####
RMSPD_gollob_1=(press_gollob_1/(g*e) + (nrep-1)*(QMRes/nrep)/nrep)##(1/2);
R_ammi_gollob_1=QMRes/(nrep*(RMSPD_gollob_1##2) - QMRES);
print RMSPD_gollob_1 R_ammi_gollob_1;
RMSPD_gollob_2=(press_gollob_2/(g*e) +(nrep-1)*(QMRes/nrep)/nrep)##(1/2);
R_ammi_gollob_2=QMRes/(nrep*(RMSPD_gollob_2##2) - QMRES);
print RMSPD_gollob_2 R_ammi_gollob_2;
RMSPD_gollob_3=(press_gollob_3/(g*e) + (nrep-1)*(QMRes/nrep)/nrep)##(1/2);
R_ammi_gollob_3=QMRes/(nrep*(RMSPD_gollob_3##2) - QMRES);
print RMSPD_gollob_3 R_ammi_gollob_3;
run;

```