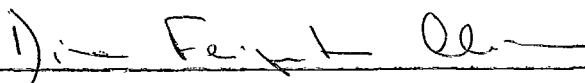


CRITÉRIOS DE ESTABILIDADE PARA O PROBLEMA
DE PREVISÃO DE MATRIZES

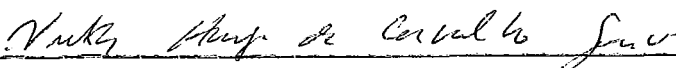
Lucia Silva Kubrusly

TESE SUBMETIDA AO CORPO DOCENTE DA COORDENAÇÃO DOS PROGRAMAS DE
PÓS-GRADUAÇÃO DE ENGENHARIA DA UNIVERSIDADE FEDERAL DO RIO DE
JANEIRO COMO PARTE DOS REQUISITOS NECESSÁRIOS PARA A OBTENÇÃO DO
GRAU DE DOUTORA EM CIÊNCIAS (D.Sc.) EM ENGENHARIA DE SISTEMAS

Aprovada por:



Prof. Dina Feigenbaum Cleiman
(Presidente)



Prof. Victor Hugo Carvalho Gouvea



Prof. Nelson Maculan Filho



Prof. Paulo Roberto de Oliveira



Prof. Frederico Antonio Azevedo de Carvalho

RIO DE JANEIRO, RJ - BRASIL
AGOSTO DE 1985

KUBRUSLY, LUCIA SILVA

Critérios de Estabilidade para o Problema de
Previsão de Matrizes. (Rio de Janeiro) 1985.

vii, 156 p. 29.7 cm (COPPE-UFRJ, D.Sc., Engenharia de Sistemas 1985).

Tese - Universidade Federal do Rio de Janeiro,
COPPE.

I. Previsão de Matrizes I. COPPE/UFRJ
II. Título (Série)

Ao meu amigo Gustavo,
porque sempre soube
ser meu amigo.

AGRADECIMENTOS

Quero agradecer a todos os meus amigos que durante os últimos quatro anos e meio me ajudaram direta ou indiretamente na elaboração desse trabalho.

Em particular agradeço à Dina, Victor, Miriam, Ricardo e Amir, porque eles foram ótimos.

Agradeço também a Gloria, Marcia, Jô, Marise e Solange que formaram a equipe responsável pela datilografia do texto.

Resumo da Tese Apresentada à COPPE/UFRJ como parte dos requisitos necessários para a obtenção do grau de Doutor em Ciências (D.Sc.)

CRITÉRIOS DE ESTABILIDADE PARA O PROBLEMA
DE PREVISÃO DE MATRIZES

Lucia Silva Kubrusly

AGOSTO 1985

Orientador: Prof^a Dina Feigenbaum Cleiman

Programa: Engenharia de Sistemas

Nesse trabalho são abordados os problemas de atualização e de previsão de matrizes. Com relação ao primeiro problema, são apresentados três tipos de métodos para sua solução: o método iterativo RAS, métodos de Programação Quadrática e métodos de Programação Linear. O problema de previsão é tratado através das "estruturas básicas" das matrizes, obtidas por métodos de agregação estatística que permitem uma considerável redução dos dados originais. Baseado na evolução dessas "estruturas básicas", são definidos alguns critérios de estabilidade para a evolução da série de matrizes.

Abstract of Thesis presented to COPPE/UFRJ as partial fulfillment of the requirements for the degree of Doctor of Science (D.Sc.)

STABILITY CRITERIA FOR THE
PROBLEM OF MATRICES PREDICTION

Lucia Silva Kubrusly

AUGUST 1985

Chairman: Prof^a Dina Feigenbaum Cleiman

Department: Systems Engineering

This work deals with the problems of updating and predicting matrices. Concerning the first problem three kinds of solution methods are discussed: iterative RAS method; Quadratic Programming methods and Linear Programming methods. The prediction problem is examined through the "basic structures" of the matrices which are determined by reduction of the original data. A statistical agregation method is used to obtain the desired reduction. Stability criteria are then defined based on the evolution of these "basic structures".

ÍNDICE

	Pág.
INTRODUÇÃO	01
CAPÍTULO I - ATUALIZAÇÃO DE MATRIZES	
I.1 - O Problema de atualização de Matrizes	04
I.2 - O Método RAS	06
I.3 - Métodos de Programação Quadrática	15
I.4 - Uma Formulação Geral para os Métodos de Atualização de Matrizes	20
I.5 - Uma Abordagem Estatística do Problema de Atualização de Matrizes	24
I.6 - Atualização por Programação Linear	28
CAPÍTULO II - PREVISÃO DE MATRIZES	
II.1 - A Previsão Termo a Termo	32
II.2 - A Previsão da Estrutura Básica da Matriz	33
II.3 - A Análise das Componentes Principais	34
CAPÍTULO III - CRITÉRIOS DE ESTABILIDADE	
III.1 - Estabilidade do Centro de Gravidade	51
III.2 - Estabilidade da Estrutura Básica da Matriz - 1º Nível de Agregação	60
III.3 - Estabilidade da Estrutura Básica da Matriz - 2º Nível de Agregação	64

III.4 - O Problema da Mudança de Ordem	71
CAPÍTULO IV - PREVISÃO E ESTUDO DOS ERROS	
IV.1 - A Previsão dos Coeficientes das Componentes Principais	79
IV.2 - A Previsão dentro das Esferas E_i	82
IV.3 - O Erro de Previsão na Esfera E_i	85
IV.4 - O Erro na Reconstrução da Matriz A_{T+1} a partir das Componentes Principais C_i^{T+1}	88
IV.5 - O Critério de Estabilidade e a Dimensão do Espaço R^q	90
CAPÍTULO V - APLICAÇÃO DOS CRITÉRIOS DE ESTABILIDADE NA PREVISÃO DA MATRIZ DO BALANÇO ENERGÉTICO	
V.1 - O Balanço Energético	103
V.2 - Os Resultados	106
CAPÍTULO VI - CONCLUSÕES	126
ANEXO	130
BIBLIOGRAFIA	153

INTRODUÇÃO

O objetivo desse trabalho é determinar critérios de estabilidade para a evolução de uma série de matrizes. As matrizes que normalmente aparecem nos problemas de previsão, são as que descrevem algum tipo de relação entre setores econômicos ou entre diversos países ou regiões. É possível citar como exemplo, a matriz de comércio exterior, a matriz do Balanço Energético, ou a matriz de insumo produto. O problema de previsão dessas matrizes apresenta normalmente duas características principais:

- i) as matrizes são de grandes dimensões,
- ii) as séries são usualmente pequenas.

O problema da dimensão da matriz pode ser contornado utilizando-se técnicas de agregação estatística que permitem uma redução dos dados originais, com uma pequena perda da informação. O objetivo dessas agregações, é manter o essencial do conjunto de dados originais (no caso, os elementos da matriz) perdendo o mínimo de informação, medida pela percentagem da variância total perdida na agregação.

Por outro lado, as séries pequenas usualmente encontradas nesses problemas, exigem métodos de previsão diversos daqueles usados para séries de grandes dimensões como por exemplo os modelos de séries temporais. A evolução das matrizes deve ser examinada, procurando-se verificar se a série de matrizes obedece a um determinado tipo de evolução no tempo, considerada estável. O presente trabalho, está dividido em seis capítulos que

serão sumariamente descritos a seguir.

O primeiro capítulo é dedicado ao tema de atualização de matrizes, que como a previsão, tem ampla utilização no problema de planejamento econômico. A diferença, é que para atualizar uma matriz, parte-se de apenas uma matriz conhecida (passada), para obter-se a matriz atual. Nesse capítulo é feita uma revisão dos métodos desenvolvidos para esse fim. São descritos o método RAS, métodos de Programação Quadrática, e métodos de Programação Linear.

No segundo capítulo é descrito o problema de previsão e a maneira pela qual será abordado, isto é, através das agregações das matrizes. O procedimento para a previsão é obter as agregações para cada matriz da série (redução dos dados), prever a agregação para o período que se deseja, e reconstruir a matriz a partir da agregação prevista. Nesse capítulo é detalhadamente descrita a técnica estatística usada na agregação de matrizes: Análise das Componentes Principais. Essa descrição detalhada é necessária, pois as propriedades dessas componentes principais extraídas das matrizes são utilizadas no desenvolvimento do estudo de estabilidade.

O terceiro capítulo é dedicado inteiramente ao estudo de estabilidade da série de matrizes. São propostos quatro critérios de estabilidade, cada um correspondendo a um tipo de evolução diferente. No primeiro critério, é verificado se existe ou não uma mudança nos elementos da matriz ao longo do tempo. O segundo procura identificar "irregularidades" na evolução das

matrizes tais como uma grande expansão em um dado tempo, ou uma expansão seguida de retração nos elementos da matriz. Os outros dois critérios são criados para as agregações das matrizes e procuram verificar "semelhanças" nas agregações ao longo do tempo.

No quarto capítulo são apresentados métodos para previsão e algumas medidas dos erros na previsão das agregações. É discutido o problema da reconstrução da matriz a partir de sua agregação, e o erro cometido nessa reconstrução é avaliado pela perda de informação devida a agregação feita.

No quinto capítulo é feita uma aplicação dos critérios de estabilidade na previsão da matriz do Balanço Energético Nacional para o ano de 1981 a partir das matrizes relativas aos anos de 1976 a 1980. A aplicação desses critérios permite analisar a evolução das agregações, possibilitando também identificar "focos de instabilidades" na evolução da série de matrizes.

Finalmente, no sexto capítulo estão as principais conclusões tiradas do estudo teórico da estabilidade na evolução de uma série de matrizes, e também do uso dos critérios definidos nesse estudo no procedimento para a previsão, como indicadores que permitem a descrição, até certo ponto, da própria evolução da série no tempo.

CAPÍTULO I

ATUALIZAÇÃO DE MATRIZES

Um dos problemas de planejamento econômico é o que se refere a quantificação das relações entre setores da economia. Para tanto é necessário dispor de instrumentos que permitam a determinação dessas relações. As matrizes de insumo-produto, de balanço energético, do comércio exterior, de bens de capital entre outras, são exemplos de instrumentos que fornecem medidas dessas relações. Ocorre porém, que na maioria dos casos, a elaboração dessas matrizes é de difícil execução além de ser muito dispendiosa, de tal modo que elas são elaboradas de tempos em tempos (com intervalos de 5 a 10 anos). Surge portanto a necessidade de atualizar da melhor maneira possível esses instrumentos de planejamento econômico. Esse capítulo é destinado a apresentação dos principais métodos de atualização de matrizes.

I.1. O Problema de Atualização de Matrizes.

Considere uma matriz $A(n \times m)$ cujos coeficientes sejam medidas de uma certa relação entre setores econômicos. Por exemplo, suponha que cada coeficiente a_{ij} forneça a quantidade de energia gasta no setor i , proveniente da fonte de energia j . Neste caso, a soma dos coeficientes por linha e por coluna.

$$\sum_j a_{ij} = L_i$$

$$\sum_i a_{ij} = C_j$$

fornecem respectivamente a quantidade total de energia gasta no setor i (L_i), e a quantidade total de energia proveniente da fonte j (C_j). Os valores L_i ($i = 1, 2, \dots, n$) e C_j ($j = 1, 2, \dots, m$) são chamados de margens da matriz A , e por corresponderem a grandezas globais, são usualmente mais fáceis de se prever que os coeficientes a_{ij} isoladamente. Essa característica é explorada na modelagem do problema de atualização de matrizes, onde é feita a hipótese de que para um certo período (atual ou futuro) não se conhece a matriz A , mas são conhecidas suas margens L_i e C_j . A formalização do problema é dada a seguir.

Sejam:

A_0 ($n \times m$) a matriz no tempo t_0

A ($n \times m$) a matriz no tempo $t > t_0$

Determine a matriz $A(n \times m)$, a partir de $A_0(n \times m)$

e de

$$L_i = \sum_j a_{ij} \quad i = 1, 2, \dots, n$$

$$C_j = \sum_i a_{ij} \quad j = 1, 2, \dots, m$$

A obtenção exata da matriz A não é possível a partir desses dados, mas pode-se obter diversas estimativas segundo diferentes métodos. Alguns desses métodos serão apresentados a seguir, a começar pelo método RAS de atualização de matrizes.

I.2. O Método RAS.

O método RAS (ver BACHARACH - 1965) de atualização de matrizes é formulado da seguinte maneira:

Determine: r_i ($i = 1, 2, \dots, n$) e s_j ($j = 1, 2, \dots, m$)

tal que:

$$a_{ij} = r_i a_{ij}^0 s_j$$

$$\sum_j a_{ij} = L_i$$

$$\sum_i a_{ij} = C_j$$

onde:

a_{ij} são os coeficientes da matriz atualizada A (que se deseja);

a_{ij}^0 são os coeficientes da matriz A_0 (conhecida);

L_i é a soma nas linhas de A (conhecida);

C_j é a soma nas colunas de A (conhecida);

$r \in R^n$ e $s \in R^m$ são os vetores multiplicativos que representam as alterações na matriz inicial.

A determinação dos vetores multiplicativos $r \in R^n$, $s \in R^m$ e dos coeficientes a_{ij} pode ser feita pelo seguinte método iterativo que ajusta alternadamente as linhas e as colunas da matriz:

Sejam L_i ($i = 1, 2, \dots, n$), C_j ($j = 1, 2, \dots, m$)
e $A_0 (n \times m) = (a_{ij}^0)$, dados.

Passo (0): $K = 0$, $a_{ij}^k = a_{ij}^0$. Se $a_i^k = \sum_j a_{ij}^k = L_i$ e

$$a_j^k = \sum_i a_{ij}^k = C_j, \text{ faça } A = A^k \text{ pare.}$$

Senão faça $k = k + 1$

Passo (1): $a_{ij}^k = a_{ij}^{k-1} \frac{L_i}{a_i^{k-1}}$ (ajusta as margens das linhas L_i)

se $a_j^k = \sum_i a_{ij}^k = C_j$, faça $A = A^k$, pare.

Senão faça $k = k + 1$

$$a_{ij}^k = a_{ij}^{k-1} \frac{C_j}{a_j^{k-1}} \text{ (ajusta as margens das colunas } C_j)$$

se $a_i^k = \sum_j a_{ij}^k = L_i$, faça $A = A^k$, pare.

Senão volte para (1)

Assim os vetores multiplicativos serão dados por

$$r_i = \prod_{k=0}^N \frac{L_i}{a_i^{2k+1}} \quad \text{e} \quad s_j = \prod_{k=0}^N \frac{C_j}{a_j^{2k+2}}$$

onde $N = \frac{\text{n}^\circ \text{ iterações}}{2} = \text{n}^\circ \text{ de ajuste por linhas e colunas}$

Note que esse algoritmo ajusta as margens das linhas L_i e posteriormente ajusta as margens das colunas C_j até que se obtenha os dois ajustes a menos de uma tolerância dada. Num estudo detalhado desse método, BACHARACH(1965) verifica a sua convergência, bem como a propriedade de conservação dos coeficientes nulos.

Aplicação do Método RAS à Matriz de Insumo Produto.

i) O Modelo de Insumo Produto

Entre as matrizes que descrevem relações entre diversos setores econômicos, destaca-se como principal instrumento de planejamento, a matriz de insumo produto definida por LEONTIEF (1941) no seu modelo de insumo produto cujo objetivo é de descrever quantitativamente as relações entre as indústrias numa economia em um dado período. Será apresentada a seguir a versão mais simples desse modelo, baseada nas seguintes hipóteses:

- i) A economia é composta de n indústrias produzindo n diferentes produtos utilizados como insumo no processo produtivo.
- ii) O processo produtivo apresenta retorno constante em escala.

iii) q_{ij} é a quantidade física do insumo i consumida pela indústria j .

iv) p_i é o preço do insumo i .

v) q_i é a quantidade total do insumo i consumida pela economia

A distribuição de cada um dos n insumos entre as n indústrias é descrito pelo seguinte sistema linear:

$$q_{11} + q_{12} + \dots + q_{1n} = q_1$$

$$q_{21} + q_{22} + \dots + q_{2n} = q_2$$

$$\begin{array}{c} \cdot \\ \cdot \\ \cdot \end{array} \quad \begin{array}{c} \cdot \\ \cdot \\ \cdot \end{array}$$

$$q_{n1} + q_{n2} + \dots + q_{nn} = q_n$$

Multiplicando-se cada equação pelo preço p_i do insumo i , obtém-se os gastos devido a cada insumo i efetuados por cada indústria j .

$$q_{11}p_1 + q_{12}p_1 + \dots + q_{1n}p_1 = q_1p_1$$

$$q_{21}p_2 + q_{22}p_2 + \dots + q_{2n}p_2 = q_2p_2$$

$$\begin{array}{c} \cdot \\ \cdot \\ \cdot \end{array} \quad \begin{array}{c} \cdot \\ \cdot \\ \cdot \end{array}$$

$$q_{n1}p_n + q_{n2}p_n + \dots + q_{nn}p_n = q_np_n$$

Por outro lado cada indústria j tem um gasto com insumos dado por

$$q_{11}p_1 + q_{21}p_2 + \dots + q_{n1}p_n = q_1p_1$$

$$q_{12}p_1 + q_{22}p_2 + \dots + q_{n2}p_n = q_2p_2$$

$$\vdots \qquad \qquad \qquad \vdots$$

$$q_{1n}p_1 + q_{2n}p_2 + \dots + q_{nn}p_n = q_np_n$$

As equações desse sistema indicam que cada indústria gasta na produção exatamente o valor da sua própria produção. Isto é, está sendo considerada a hipótese de que não existe excedente econômico, e portanto toda a produção é utilizada no próprio processo produtivo.

Definindo-se como coeficiente técnico de produção a quantidade de insumos i necessária para a fabricação de uma unidade de produto j , isto é:

$$a_{ij} = \frac{q_{ij}}{q_j}$$

pode-se reescrever os dois sistemas acima:

1º Sistema: quantidades físicas.

$$a_{11}q_1 + a_{12}q_2 + \dots + a_{1n}q_n = q_1$$

$$\begin{array}{r}
 a_{21}q_1 + a_{22}q_2 + \dots + a_{2n}q_n = q_2 \\
 \vdots \\
 a_{n1}q_1 + a_{n2}q_2 + \dots + a_{nn}q_n = q_n
 \end{array} \tag{I.1}$$

2º Sistema: Preços

$$\begin{array}{r}
 a_{11}p_1 + a_{21}p_2 + \dots + a_{n1}p_n = p_1 \\
 a_{12}p_1 + a_{22}p_2 + \dots + a_{n2}p_n = p_2 \\
 \vdots \\
 a_{1n}p_1 + a_{2n}p_2 + \dots + a_{nn}p_n = p_n
 \end{array} \tag{I.2}$$

Considerando a matriz $A(n \times n)$ formada pelos coeficientes técnicos de produção e os vetores q e $p \in \mathbb{R}^n$ que fornecem respectivamente as quantidades físicas e os preços das n mercadorias produzidas, os dois sistemas podem ser escritos:

$$Aq = q$$

$$A'p = p$$

onde A' é a matriz transposta de A .

A matriz A que descreve o processo produtivo através dos coeficientes técnicos de produção a_{ij} é chamada de

matriz de insumo-produto e através dela é possível analisar as relações de produção econômica. Embora no modelo apresentado a matriz A seja quadrada, não é necessária essa hipótese para a utilização do método RAS para sua atualização, conforme será visto a seguir.

ii) O Método RAS e a Matriz de Insumo Produto

Se $A(n \times m)$ é a matriz de insumo produto, não é usual que se conheça diretamente a soma das linhas e colunas de A (necessárias para a utilização do Método RAS), mas sim as somas das linhas e colunas da matriz de fluxo de mercadoria $X(n \times m)$ (correspondentes as quantidades físicas totais). A relação entre X e A é dada por:

$$X = A Q$$

onde $Q(m \times m)$ é uma matriz diagonal que fornece a produção total para cada uma das indústrias do sistema econômico em questão, isto é:

$$Q = \text{diag}(q)$$

Segundo o método RAS, a matriz atualizada será:

$$A = R A_0 S$$

onde $R(n \times n)$ e $S(m \times m)$ são matrizes diagonais formadas pelos vetores r e s .

Supondo conhecidos A_0 , Q e as margens da matriz de fluxo X ($u_i = \sum_j x_{ij}$, $v_j = \sum_i x_{ij}$) tem-se:

$$X = A Q$$

$$u_i = \sum_j x_{ij}$$

$$v_j = \sum_i x_{ij}$$

Sabendo-se que

$$A = R A_0 S$$

Então pode-se escrever

$$X = R A_0 S Q$$

$$X = R A_0 Q S$$

porque Q e S sendo diagonais $SQ = QS$. A matriz $Y = A_0 Q$ é conhecida porque A_0 e Q são supostas conhecidas, portanto o problema de se obter a matriz X de fluxos pode ser formulado segundo o método RAS de atualização.

Determine $r \in R^n$ e $s \in R^m$ tal que

$$x_{ij} = r_i y_{ij} s_j$$

$$\text{s.a: } \sum_j x_{ij} = u_i$$

$$\sum_i x_{ij} = v_j$$

A matriz dos coeficientes técnicos será obtida por:

$$A = X Q^{-1}$$

O Método RAS Modificado.

O método RAS de atualização de matrizes pode ser ligeiramente modificado para atender aos casos em que além das margens atuais, são também considerados conhecidos* alguns dos seus atuais coeficientes. Esse método é conhecido como RAS modificado e operacionalmente é obtido, descontando-se nos totais das linhas e colunas o valor desses coeficientes que receberão valor zero provisoriamente. Procede-se então o método RAS. Como esse método preserva os zeros existentes, ao final acrescenta-se o valor conhecido para aqueles coeficientes, obtendo-se automaticamente o ajuste para as margens.

Outra maneira de se formular o problema de atualização de matrizes, é através da formulação de um problema de programação matemática. No método RAS de atualização busca-se determinar os vetores r e s que representam as alterações nos coeficientes de A_0 . Uma outra maneira de modelar o problema, é supor que de um período para o outro as alterações nos coeficientes a_{ij} são as mínimas possíveis. Isto é, são mínimas tais

(*) Trata-se do caso em que foi possível se obter boas previsões para esses coeficientes, antes de se efetuar a atualização da matriz A_0 .

que atendam as restrições para as margens da matriz A. Assim o problema de atualização pode ser escrito como (*):

Determine x_{ij} $i = 1, 2, \dots, n$, $j = 1, 2, \dots, m$ tal que

$$\min f(x_{ij})$$

$$\text{s.a: } \sum_{j=1}^m x_{ij} = L_i$$

$$\sum_{i=1}^n x_{ij} = C_j$$

$$x_{ij} \geq 0$$

A função $f(x_{ij})$ deve ser escolhida tal que descreva a alteração entre a matriz A_0 e a nova matriz.

I.3. Método de Programação Quadrática.

As formas quadráticas da função f são as mais frequentemente usadas, baseadas na idéia de ajuste estatísticos da do pela distância do χ^2 . Serão discutidas a seguir algumas dessas formulações.

(*) A fim de evitar as notações a_{ij}^0 para os coeficientes da matriz A_0 e a_{ij} para os coeficientes da matriz A, será de agora em diante adotada a seguinte notação:

x_{ij} para os coeficientes de A

a_{ij} para os coeficientes de A_0

i) Formulações de FRIEDLANDER (1961)

Determine $X = (x_{ij})$ $i = 1, 2, \dots, n$; $j = 1, 2, \dots, m$

tal que:

$$\min \sum_{ij} \frac{(x_{ij} - a_{ij})^2}{a_{ij}}$$

$$\text{s.a: } \sum_j x_{ij} = L_i$$

$$\sum_i x_{ij} = C_j$$

onde $A = (a_{ij})$, L_i e C_j são conhecidos.

Para resolver esse problema Friedlander propõe o método dos multiplicadores da Lagrange.

$$\min f = \sum_{ij} \frac{(x_{ij} - a_{ij})^2}{a_{ij}} = \min \sum_{ij} \frac{(x_{ij} - a_{ij})^2}{2a_{ij}}$$

$$\text{s.a: } L_i - \sum_j x_{ij} = 0 \quad i = 1, 2, \dots, n$$

$$C_j - \sum_i x_{ij} = 0 \quad j = 1, 2, \dots, m$$

A função lagrangeana será:

$$L = \sum_{ij} \frac{(x_{ij} - a_{ij})^2}{2a_{ij}} + \sum_i \lambda_i (L_i - \sum_j x_{ij}) + \sum_j \mu_j (C_j - \sum_i x_{ij})$$

$$L = \sum_{ij} \frac{(x_{ij} - a_{ij})^2}{2a_{ij}} + \sum_i \lambda_i L_i - \sum_i \lambda_i \sum_j x_{ij} + \sum_j \mu_j C_j - \sum_j \mu_j \sum_i x_{ij}$$

Derivando em relação a x_{ij}

$$\frac{\partial L}{\partial x_{ij}} = \frac{(x_{ij} - a_{ij})}{a_{ij}} - \lambda_i - \mu_j = 0$$

$$\frac{x_{ij}}{a_{ij}} - 1 - \lambda_i - \mu_j = 0$$

$$x_{ij} = a_{ij} + \lambda_i a_{ij} + \mu_j a_{ij}$$

Somando para todo i :

$$\sum_i x_{ij} = \sum_i a_{ij} + \sum_i \lambda_i a_{ij} + \mu_j \sum_i a_{ij} = C_j$$

Mas se $A = (a_{ij})$ é a matriz conhecida, a soma nas linhas da matriz A será:

$$\sum_i a_{ij} = C_j^0$$

$$\text{Assim } C_j = C_j^0 + \sum_i \lambda_i a_{ij} + \mu_j C_j^0 \quad (\text{I.3})$$

Somando para todo j e chamando $\sum_j a_{ij} = L_i^0$ temos:

$$\sum_j x_{ij} = \sum_j a_{ij} + \lambda_i \sum_j a_{ij} + \sum_j \mu_j a_{ij} = L_i$$

$$L_i = L_i^0 + \lambda_i L_i^0 + \sum_j \mu_j a_{ij} \quad (I.4)$$

Podemos escrever as equações (I.3) e (I.4).

$$C = C^0 + A^1 + MC^0 \quad (I.5)$$

$$L = L^0 + \Lambda L^0 + A\mu \quad (I.6)$$

onde $C, C^0, \mu \in R^m$

$L, L^0, \lambda \in R^n$

$M (m \times m) = \text{diag} (\mu_j)$

$\Lambda (n \times n) = \text{diag} (\lambda_i)$

$A (n \times m) = \text{matriz inicial conhecida.}$

As equações (I.5) e (I.6) formam um sistema de $m + n$ equações cuja solução fornece os $(m + n)$ multiplicadores de Lagrange.

No seu trabalho Friedlander faz ainda uma relação entre essa solução e a solução obtida por um método iterativo de incrementos sucessivos com relação aos dados originais (um método semelhante ao método RAS). Essa tentativa de relacionar esses dois tipos de métodos (iterativo buscando uma solu

ção viável, e método de Programação Quadrática] será visto posteriormente com mais detalhes.

ii) Formulação de BACHEM e KORTE (1980).

Nessa formulação a função quadrática a ser minimizada é:

$$f = \sum_{ij} \frac{(x_{ij} - a_{ij})^2}{2a_{ij}^2}$$

$$\text{s.a: } L_i = \sum_j x_{ij}$$

$$C_j = \sum_i x_{ij}$$

fornecendo uma solução para x_{ij} dependente dos multiplicadores de Lagrange λ_i e μ_j :

$$x_{ij} = a_{ij} + \lambda_i a_{ij}^2 + a_{ij}^2 \mu_j$$

iii) Formulação de BACHARACH (1970)

Nessa formulação a função distância é utilizada como critério para o ajuste:

$$\min f = \sum_{ij} (x_{ij} - a_{ij})^2$$

$$\text{s.a: } \sum_j x_{ij} = L_i$$

$$\sum_i x_{ij} = C_j$$

cuja solução para x_{ij} em função dos multiplicadores de Lagrange é:

$$x_{ij} = a_{ij} + \lambda_i + \mu_j$$

Isto é, a cada elemento a_{ij} acrescenta-se uma parcela relativa a i -ésima linha e outra relativa a j -ésima coluna.

I.4. Uma Formulação Geral para os Métodos de Atualização de Matrizes.

Para todos os métodos de atualização de matrizes até agora vistos, chega-se a uma equação para x_{ij} dependendo de dois tipos de multiplicadores (por linhas e por colunas) e do valor conhecido a_{ij} . O problema fica determinado quando a esta equação para x_{ij} , acrescenta-se as equações de vínculo

$$\sum_i x_{ij} = C_j$$

$$\sum_j x_{ij} = L_i$$

BACHEM e KORTE (1981) propõe uma relação geral entre x_{ij} e a_{ij} baseados na idéia que qualquer mudança dos coeficientes da matriz A para a matriz X devidamente ponderada, é

proporcional a uma combinação linear entre a média aritmética e o quadrado da média geométrica das mudanças por linhas e colunas λ_i e μ_j , isto é:

$$\frac{x_{ij} - a_{ij}}{w_{ij}} = \frac{(\lambda_i + \mu_j)}{2} + \tau(\sqrt{\lambda_i \mu_j}) \quad 0 \leq \tau \leq 1$$

Considerando X ($n \times m$)

A ($n \times m$)

$\lambda \in R^n$ e $\Lambda = \text{diag}(\lambda)$

$\mu \in R^m$ e $M = \text{diag}(\mu)$

a relação geral para X e A é dada por:

$$X = A + \Lambda W + W M + \tau \Lambda W M \quad (I.7)$$

onde W é a matriz dos pesos.

Voltando rapidamente aos métodos anteriormente abordados, as relações entre X e A são dadas por:

i) Método RAS

$$x_{ij} = r_i a_{ij} s_j$$

ii) Friedlander

$$x_{ij} = a_{ij} + \lambda_i a_{ij} + \mu_j a_{ij}$$

iii) Bachem e Korte

$$x_{ij} = a_{ij} + \lambda_i a_{ij}^2 + a_{ij}^2 \mu_j$$

iv) Bacharach

$$x_{ij} = a_{ij} + \lambda_i + \mu_j$$

Definindo-se valores convenientes para W , τ , λ e μ é possível escrever todos esses métodos como derivados da relação geral (I - 7).

i) Para o método RAS, define-se

$$W = A, \quad \tau = 1, \quad \lambda = r - e_n, \quad \mu = s - e_m$$

onde $e_n \in R^n$ e $e_m \in R^m$ são vetores formados de 1's.

$$X = A + (RA - A) + AS - A + (RA - A)(S - I)$$

$$X = RA + AS - A + RAS - RA - AS + A = RAS$$

$$\text{ou } x_{ij} = r_i a_{ij} s_j$$

ii) Friedlander

$$W = A, \quad \tau = \rho, \quad \lambda \text{ e } \mu \text{ os múltiplos de Lagrange}$$

$$X = A + \Lambda A + AM$$

$$\text{ou } x_{ij} = a_{ij} + \lambda_i a_{ij} + a_{ij} \mu_j$$

iii) Bachem e Korte

$$W_{ij} = a_{ij}^2, \tau = 0, \lambda \text{ e } \mu \text{ os multiplicadores de Langrange}$$

$$W = \begin{vmatrix} a_{11}^2 & \dots & a_{1m}^2 \\ \vdots & & \vdots \\ a_{n1}^2 & \dots & a_{nm}^2 \end{vmatrix}$$

$$X = A + \Lambda W + WM$$

$$\text{ou } x_{ij} = a_{ij} + \lambda_i a_{ij}^2 + a_{ij}^2 \mu_j$$

iv) Bacharach

$$W = \begin{vmatrix} 1 & \dots & 1 \\ \vdots & & \vdots \\ 1 & \dots & 1 \end{vmatrix}, \tau = 0, \lambda \text{ e } \mu \text{ os multiplicadores de Langrange}$$

$$X = A + \begin{vmatrix} \lambda_1 & \dots & \lambda_1 \\ \vdots & & \vdots \\ \lambda_n & \dots & \lambda_n \end{vmatrix} + \begin{vmatrix} \mu_1 & \dots & \mu_m \\ \vdots & & \vdots \\ \mu_1 & \dots & \mu_m \end{vmatrix}$$

$$\text{ou } x_{ij} = a_{ij} \cdot \lambda_i + \mu_j$$

I.5. Uma Abordagem Estatística do Problema de Atualização de Matrizes.

O problema de atualização de matrizes pode ser visto como um problema de estimação estatística no qual os coeficientes da matriz X são considerados como variáveis aleatórias cujos valores deseja-se estimar. Stone (1942) e posteriormente Byron (1978) deram um tratamento estatístico para esse problema, no caso particular em que as matrizes A e X são quadradas. Segundo esse enfoque, determina-se uma estimativa para a nova matriz tal que minimize uma função de perda quadrática. Essa função é obtida considerando-se os elementos da matriz $X = (x_{ij})$ alinhados formando um vetor $X \in R^{m^2}$, e a matriz $V(m^2 \times m^2)$ que forneça as estimativas das covariâncias entre os elementos de X , e os desvios entre os valores iniciais e os valores estimados. Assim a função perda a ser minimizada é definida por:

$$\min Z = \frac{1}{2} d' V^{-1} d + \lambda' (G x - h)$$

onde:

$d = (x - a) \in R^{m^2}$ é o vetor dos desvios entre os valores estimados e os valores iniciais.

$V(m^2 \times m^2)$ é a estimativa da matriz de covariância

$\lambda \in \mathbb{R}^{m+m}$ é o vetor dos multiplicadores de Lagrange relativos aos $(m + m)$ vínculos.

$h \in \mathbb{R}^{m+m}$ são os valores dados para as somas das linhas e das colunas da matriz X . Isto é os vínculos

$$\sum_j x_{ij} = L_i \quad i = 1, \dots, m$$

$$\sum_i x_{ij} = C_j \quad j = 1, \dots, m$$

São representados por:

$$G x = h \quad \text{onde}$$

$$G = \begin{pmatrix} 1 & \dots & 1 & 0 & \dots & 0 & 0 & \dots & 0 & \dots & 0 \\ 0 & \dots & 0 & 1 & \dots & 1 & 0 & \dots & 0 & \dots & 0 \\ \vdots & & & & & & \vdots & & & & \vdots \\ \vdots & & & & & & \vdots & & & & \vdots \\ 0 & \dots & 0 & 0 & \dots & 0 & 0 & \dots & 0 & \dots & 1 \dots 1 \\ 10 \dots 010 & \dots & 0 & 1 & \dots & 0 & \dots & 10 \dots 0 \\ 01 \dots 001 & \dots & 0 & 0 & 1 & \dots & 0 & \dots & 01 \dots 0 \\ \vdots & & & \vdots & & & \vdots & & & & \vdots \\ \vdots & & & \vdots & & & \vdots & & & & \vdots \\ 00 \dots 100 & \dots & 1 & 0 & 0 & \dots & 1 & \dots & 00 \dots 1 \end{pmatrix}$$

$$h = [L_1, \dots, L_m, C_1, \dots, C_m]^T$$

A fim de interpretar melhor a função Z , considere o caso particular em que V é diagonal. Nesse caso a função perda será:

$$Z = \sum_{k=1}^{m^2} \frac{(d_k)^2}{\sigma_k^2} - \lambda' (Gx - h)$$

onde σ_k^2 é a variância de x_k .

Assim os desvios são ponderados pela variância do coeficiente correspondente. Uma maior variância de x_k permite um desvio d_k maior, ao passo que se x_k tiver uma pequena variância, exige-se que a estimativa para x_k seja mais próxima do valor inicial a_k . Isto é, considera-se o desvio causado pelos seguintes fatores:

- i) a mudança da variável x_k ocorrida em um período
- ii) a dispersão da variável x_k .

A ponderação pelas variâncias, tem como objetivo "descontar" o fator dispersão do desvio ocorrido.

No caso geral onde V não é necessariamente diagonal, a minimização da função

$$Z = \frac{1}{2} d' V^{-1} d + \lambda' (Gx - h) = \frac{1}{2} (x - a)' V^{-1} (x - a) + \lambda' (Gx - h)$$

é obtida derivando-se Z em relação X e igualando a zero:

$$\frac{\partial Z}{\partial x} = V^{-1} (x - a) + G' \lambda = 0$$

$$V^{-1} (x - a) = - G' \lambda$$

$$V^{-1} x = V^{-1} a - G' \lambda$$

$$X = a - V G' \lambda \tag{I.8}$$

Por outro lado, das equações de vínculos

$$Gx = h \tag{I.9}$$

Das equações (I.8) e (I.9)

$$V G' \lambda = a - x$$

$$G V G' \lambda = Ga - h$$

$$\lambda = (G V G')^{-1} (Ga - h)$$

$$\lambda = (G V G')^{-1} (Ga - h)$$

Portanto, de (I.8)

$$x = a - V G' [G V G']^{-1} (G_a - h)$$

Embora essa solução para o problema de atualização tenha o inconveniente de necessitar da estimativa da matriz de covariância, por outro lado, o estimador de x assim determinado é um estimador BLUE (melhor estimador linear não tendencioso) conforme foi demonstrado por Theil (1961). Portanto, embora a função a ser minimizada não seja muito diferente das funções quadráticas definidas anteriormente, essa abordagem estatística do problema, ao contrário das demais, permite uma avaliação do estimador obtido.

I.6. Atualização por Programação Linear.

Embora os métodos RAS e de programação quadrática sejam os mais frequentemente usados no problema de atualização de matrizes, algumas tentativas de modelar esse problema foram feitas utilizando-se programação linear: MATUZEWSKI, PITTS e SAWYER (1964) e mais tarde NIJKAMP e PAELINK (1974) foram alguns dos que propuseram uma modelagem desse tipo. A idéia básica ainda é minimizar os desvios entre os valores conhecidos e os valores a serem estimados, diferindo da formulação quadrática apenas pelo fato de que nesse caso, são considerados os desvios absolutos. O problema é formulado da seguinte maneira.

Determine a matriz $X = (x_{ij})$ tal que

$$\min \sum_{ij} |x_{ij} - a_{ij}| C_{ij}, \text{ com } C_{ij} = \frac{1}{a_{ij}}$$

$$\text{s.a: } \sum_i x_{ij} = C_j$$

$$\sum_j x_{ij} = L_i$$

$$1/2 \leq x_{ij}/a_{ij} \leq 2$$

A terceira restrição adicionada impõe limites para a variação dos coeficientes, visando evitar um número muito grande de surgimento de zeros para os coeficientes da matriz X , compensado por valores comparativamente muito grandes para os demais valores. Esses limites são na verdade bastante difíceis de se estabelecer sendo uma das principais críticas feitas a essa formulação.

Supondo que os coeficientes da matriz X ou sofrem um acréscimo (x^+) ou sofrem um decréscimo (x^-) ou ainda não se alteram em relação ao coeficiente inicial, pode-se escrever:

$$x_{ij} = a_{ij} + x_{ij}^+ - x_{ij}^-$$

$$x_{ij}^+ \geq 0 \quad x_{ij}^- \geq 0$$

sendo que para cada par (ij) , no máximo uma das duas variáveis (x_{ij}^+ ou x_{ij}^-) é diferente de zero. Os desvios absolutos podem ser escritos da seguinte maneira:

$$|x_{ij} - a_{ij}| = |a_{ij} + x_{ij}^+ - x_{ij}^- + a_{ij}| = |x_{ij}^+ - x_{ij}^-|$$

$$= (x_{ij}^+ + x_{ij}^-)$$

Portanto o problema pode ser reescrito como:

$$\min \sum_{ij} (x_{ij}^+ + x_{ij}^-) C_{ij}$$

$$\text{s.a: } \sum_i (a_{ij} + x_{ij}^+ - x_{ij}^-) = C_j$$

$$\sum_j (a_{ij} + x_{ij}^+ - x_{ij}^-) = L_i$$

$$1/2 \leq (a_{ij} + x_{ij}^+ - x_{ij}^-)/a_{ij} \leq 2$$

$$x_{ij}^+ \geq 0 \quad x_{ij}^- \geq 0$$

ou ainda:

$$\min \sum_{ij} (x_{ij}^+ + x_{ij}^-) C_{ij}$$

$$\text{s.a: } \sum_i (x_{ij}^+ - x_{ij}^-) = C_j - \sum_i a_{ij}$$

$$\sum_j (x_{ij}^+ - x_{ij}^-) = L_i - \sum_j a_{ij}$$

$$0 \leq x_{ij}^+ \leq a_{ij}$$

$$0 \leq x_{ij}^- \leq a_{ij}/2$$

Matuzewski, Pitts e Sawyer consideram esse problema como um problema de transporte onde os limites superiores das variáveis x^+ e x^- são vistos como restrições de capacidade.

Os métodos de atualização usando programação linear são, segundo algumas análises comparativas já realizadas (ver Teixeira - 1978) inferiores aos métodos RAS e de programação quadrática. Nessas análises é usada como medida de comparação, o erro cometido na atualização através desses métodos.

CAPITULO II

PREVISÃO DE MATRIZES

O problema de atualização de matrizes difere do problema de previsão porque no segundo caso busca-se obter uma estimativa para a matriz X levando-se em conta não apenas uma matriz inicial A , mas uma série de matrizes no tempo A_1, A_2, \dots, A_T . Isto é, no caso da previsão utiliza-se a informação sobre a tendência de mudança apresentada nos últimos períodos para os coeficientes a_{ij} .

II.1 - A Previsão Termo a Termo

Van der Ploeg (1982) tratou do problema de previsão de matrizes como um problema de se ajustar uma série de matrizes econômicas supondo que cada uma delas contém erros de diversas naturezas. A matriz é representada por um vetor $\in \mathbb{R}^{nm}$ tal que o vetor observado no período t (x_t^{obs}) contém uma parcela devida a medida de ruído aleatório (ϵ_t^x), outra devida a um erro sistemático que ocorre em todos os períodos (ϵ^{sist}), e outra contendo erros proporcionais (ϵ^{prop}) a uma série de dados conhecida (Y_t), ou seja

$$x_t^{obs} = x_t + \epsilon_t^x + \epsilon_m^{sist} + Y_t \epsilon^{prop}$$

A estimativa para x_t^{obs} deve ser obtida de tal modo que obedeça as equações de vínculos para as margens, e minimize o vetor dos erros:

$$\epsilon' = (\epsilon^x, \epsilon^{sist}, \epsilon^{prop})$$

No seu trabalho Van der Ploeg analisa as hipóteses de autocorrelação dos erros uniforme, e de erro sistemático com tendência cíclica, e aplica seu modelo na matriz agregada (de produções globais) para a economia inglesa. Naturalmente a aplicação seria extremamente difícil para a matriz desagregada (matriz de insumo-produto) devido entre outras coisas, às suas grandes dimensões.

II.2 - A Previsão da Estrutura Básica da Matriz

Como vimos anteriormente a previsão de matrizes apresenta dois problemas principais. Primeiro as matrizes econômicas são usualmente de grandes dimensões, tornando a previsão de todos os seus coeficientes um problema excessivamente trabalhoso. Outro aspecto de ordem prática, é que na maioria das vezes não se tem uma série de matrizes suficientemente grande para que se possa fazer estimativa de erros e tendências confiáveis. Isto é, a previsão usando modelos de séries temporais, não são aplicáveis na maioria das vezes devida a pequena dimensão da série de dados disponíveis.

A fim de contornar o problema da dimensão das matrizes, ao invés de se trabalhar com todos os coeficientes de cada matriz, passa-se a trabalhar com agregações dessas matrizes. Essas agregações devem ser tais que contenham o máximo de informações sobre a matriz com um mínimo de elementos. Em outras palavras, deve-se encontrar o essencial da matriz, e por isso mesmo essas agregações são denominadas de estrutura básica da matriz.

Uma maneira de se determinar essa estrutura básica é através da utilização de técnicas estatísticas de agregação (veja STEMMELEN (1977) e LE FOLL, BURTSCHY (1983)), das quais se destaca a Análise das Componentes Principais. Na utilização dessa técnica, que será descrita com detalhes a seguir, cada matriz A_t ($n \times m$) da série é vista como uma matriz de dados, considerando-se cada coluna x_k , $k = 1, 2, \dots, m$ (ou cada linha x_ℓ , $\ell = 1, 2, \dots, m$) como uma variável aleatória da qual foram feitas n observações $a_{\ell k}$ (k fixo, $\ell = 1, 2, \dots, n$). Na verdade para cada matriz A_t serão extraídas duas agregações: uma relativa às colunas e outra relativa às linhas. Esse par de agregações forma a chamada estrutura básica da matriz. Como essas agregações são realizadas pela mesma técnica, segundo os mesmos objetivos, muitas vezes será feita neste trabalho referência a apenas uma delas, mas todos os resultados obtidos para a agregação das colunas é válido para a agregação das linhas.

II.3 - A Análise das Componentes Principais

- Agregação das colunas

Na matriz A ($n \times m$) considere suas colunas X_k ($k = 1, 2, \dots, m$) como variáveis aleatórias sobre as quais são feitas n observações $a_{\ell k}$ (k fixo, $\ell = 1, 2, \dots, n$). Tem-se portanto um conjunto de m variáveis sobre a qual são feitas n observações. A Análise das Componentes Principais tem como objetivo descrever a configuração dessas observações no espaço das variáveis. Na figura (II.1) é apresentado um exemplo em que para duas variáveis X_1 e X_2 são feitas sete observações.

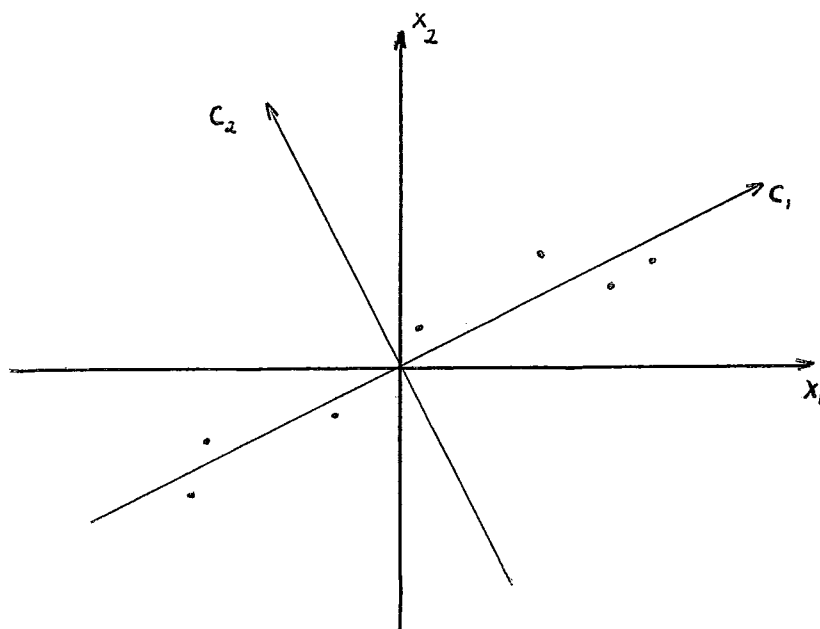


Figura II.1 - As Componentes Principais de um Conjunto de Observações

As componentes principais C_1 e C_2 desse conjunto de pontos acha-se representada na figura, e fornece as direções de maior dispersão dos pontos observados. Na verdade, trata-se apenas de uma rotação ortogonal do sistema de referência original (dado pelas variáveis X_1 e X_2), onde C_1 é uma combinação linear de X_1 e X_2 , na direção de maior dispersão dos pontos. C_2 é também uma combinação linear de X_1 e X_2 ortogonal a C_1 .

Voltando a matriz A ($n \times m$) representando n observações de m variáveis, pode-se considerar a nuvem de n pontos no espaço R^m . As componentes principais extraídas serão combinações lineares das m variáveis, tais que C_1 tenha variância máxima; C_2 tenha variância máxima e seja ortogonal a C_1 (em termos estatísticos isto equivale a não correlação entre as componentes); assim sucessivamente até que C_m será uma combinação linear das m variáveis com variância máxima, e ortogonal a (ou não correlacionada com) $C_1, C_2 \dots C_{m-1}$.

Pode-se escrever:

$$\begin{aligned}
 C_1 &= \sum_{k=1}^m c_{k1} X_k && \text{tal que } \text{var}(C_1) = \max \\
 C_2 &= \sum_{k=1}^m c_{k2} X_k && \text{tal que } \text{var}(C_2) = \max \\
 & && \text{corr}(C_1, C_2) = 0 \\
 & \vdots && \vdots \\
 & \vdots && \vdots \\
 C_m &= \sum_{k=1}^m c_{km} X_k && \text{tal que } \text{var}(C_m) = \max \\
 & && \text{corr}(C_i, C_m) = 0 \quad i=1,2, \dots, m-1
 \end{aligned}$$

Na extração dessas componentes C_i ($i=1,2, \dots, m$) ocorre que as variâncias são decrescentes devido a restrição crescente de não correlação entre elas. Usualmente as últimas componentes extraídas, têm variâncias tão pequenas que podem ser desprezadas. (Na figura II.1), a dispersão na direção de C_1 é tão maior que na direção C_2 , que a configuração dos pontos pode ser quase totalmente descrita na direção de C_1). A redução na dimensão inicial do problema é obtida mantendo-se apenas as p primeiras componentes, as quais descrevem uma grande percentagem da variância total do conjunto de m variáveis.

- Determinação das Componentes Principais (ver JOHNSON, WICHERN - 1982)

Voltando-se a primeira componente C_1

$$C_1 = c_{11} X_1 + c_{21} X_2 + \dots + c_{m1} X_m$$

Considerando as variáveis $X = (X_1, \dots, X_m)'$ e o vetor dos coeficientes $c_1 = (c_{11}, c_{21}, \dots, c_{m1})'$ pode-se escrever:

$$C_1 = c_1' X$$

Dessa forma o problema de determinar C_1 passa a ser o de determinar o vetor $c_1 \in R^m$ cuja direção seja a da maior variância do conjunto de pontos. Já que o importante é a direção dada pelo vetor c_1 , pode-se considerar sua norma unitária, isto é,

$$c_1' c_1 = 1$$

Sendo

$$C_1 = c_1' X,$$

a variância de C_1 será:

$$\text{var}(C_1) = c_1' \Sigma c_1$$

onde Σ é a matriz das covariâncias de X . Assim o problema a ser resolvido é:

Determine $c_1 \in R^m$ tal que

$$c_1' \Sigma c_1 = \max$$

$$\text{s.a: } c_1' c_1 = 1$$

A solução desse problema pode ser obtida utilizando-se os multiplicadores de Lagrange. A função lagrangeana será:

$$L = c_1' \Sigma c_1 - \lambda_1 (c_1' c_1 - 1)$$

Derivando em relação a c_1 e igualando a zero:

$$2 (\Sigma - \lambda_1 I) c_1 = 0$$

$$\Sigma c_1 = \lambda_1 c_1$$

Esta é a equação característica da matriz Σ para extrair seus auto valores λ_1 e auto vetores c_1 . Por outro lado:

$$c_1' \Sigma c_1 = c_1' \lambda_1 c_1$$

como $c_1' c_1 = 1$

$$c_1' \Sigma c_1 = \lambda_1 = \text{var}(C_1)$$

Assim a variância de C_1 é um auto valor da matriz Σ . Como essa variância deve ser máxima, λ_1 é o maior auto valor de Σ . O vetor c_1 será o auto vetor correspondente normalizado.

Para se encontrar a segunda componente C_2 o procedimento é análogo, só introduzindo-se a restrição de não correlação, que em termos geométricos, equivale a restrição de ortogonalidade.

Sendo

$$C_2 = \sum_{k=1}^m c_{k2} X_k$$

$$C_2 = c_2' X$$

$$\text{var}(C_2) = c_2' \Sigma c_2$$

O problema fica da seguinte forma:

Determine $c_2 \in \mathbb{R}^m$ tal que

$$c_2' \Sigma c_2 = \text{var}(C_2) = \max$$

$$\text{s.a: } c_2' c_2 = 1$$

$$c_2' c_1 = 0$$

A última restrição garante a não correlação entre C_1 e C_2 . Da mesma maneira, usando os multiplicadores de Lagrange:

$$L = c_2' \Sigma c_2 - \lambda_2 (c_2' c_2 - 1) - \mu c_2' c_1$$

Derivando em relação a c_2 e igualando a zero:

$$2 (\Sigma - \lambda_2 I) c_2 - \mu c_1 = 0$$

Pré-multiplicando por c_1'

$$2 c_1' \Sigma c_2 - \lambda_2 c_1' c_2 - \mu c_1' c_1 = 0$$

$$2 c_1' \Sigma c_2 - \mu = 0$$

Mas do primeiro problema:

$$\Sigma c_1 - \lambda_1 c_1 = 0$$

Pré-multiplicando por c_2'

$$c_2' \Sigma c_1 - \lambda_1 c_2' c_1 = 0$$

Então $c_2' \Sigma c_1 = 0 \rightarrow \mu = 0$

Assim o vetor c_2 deve satisfazer:

$$(\Sigma - \lambda_2 I) c_2 = 0$$

$$\Sigma c_2 = \lambda_2 c_2$$

Ou seja c_2 e λ_2 são auto vetor e auto valor de Σ . E da mesma forma que no primeiro problema:

$$\Sigma c_2 = \lambda_2 c_2$$

$$c_2' \Sigma c_2 = c_2' \lambda_2 c_2$$

$$c_2' \Sigma c_2 = \lambda_2 = \text{var}(C_2)$$

Então escolhe-se λ_2 o maior possível, diferente de λ_1 . Será o segundo maior auto valor de Σ . O auto vetor c_2 correspondente fornecerá os coeficientes da componente C_2 .

O procedimento para determinar as demais componentes é o mesmo, isto é, no final são extraídos todos os auto valores e auto vetores da matriz Σ . Operacionalmente falando, determinar os componentes principais de um conjunto de dados, é extrair os auto vetores de sua matriz de covariância.

A Interpretação das Componentes Principais

Cada componente principal terá variância igual ao auto valor de mesma ordem da matriz de covariância. Se são extraídas todas as m componentes, a variância total do conjunto original de variáveis é totalmente reproduzida.

$$\lambda_1 + \lambda_2 + \dots + \lambda_m = \text{var total} = \text{tr } \Sigma \quad (\text{II-1})$$

Onde $\text{tr } \Sigma$ é a soma dos elementos da diagonal da matriz de covariância, ou seja a soma das variâncias das m variáveis.

$$\lambda_1 + \lambda_2 + \dots + \lambda_m = \text{tr } \Sigma = \sum_{i=1}^m \sigma_i^2$$

Portanto a medida da importância descritiva de uma componente C_i qualquer, é dada pela proporção da variância total, descrita por ela:

$$\frac{\lambda_i}{\text{tr } \Sigma} = \frac{\lambda_i}{\sum_{i=1}^m \lambda_i}$$

Se são mantidas $p < m$ componentes, a proporção da variância explicada será

$$\frac{\sum_{i=1}^p \lambda_i}{\sum_{i=1}^m \lambda_i} \quad (\text{II.2})$$

A combinação linear das variáveis que compõe a componente C_i

$$C_i = c_{1i} X_1 + c_{2i} X_2 + \dots + c_{mi} X_m$$

tem como coeficiente as coordenadas do i -ésimo auto vetor da matriz de covariância Σ . O sinal e a grandeza de c_{ki} indicam o sentido e a contribuição da k -ésima variável para a i -ésima componente. A covariância entre a variável X_k e a componente C_i será:

$$\text{Cov}(X_k, C_i) = E(X_k, C_i) \quad (\text{supondo que as variáveis são centradas (média zero)})$$

$$\begin{aligned} \text{Cov}(X_k, C_i) &= E \left[X_k, (c_{1i} X_1 + c_{2i} X_2 + \dots + c_{mi} X_m) \right] \\ &= E \left[c_{1i} X_1 X_k + c_{2i} X_2 X_k + \dots + c_{mi} X_m X_k \right] \end{aligned}$$

onde c_{ki} $k = 1, 2, \dots, m$ são constantes, portanto:

$$\text{Cov}(X_k, C_i) = c_{1i} E(X_1 X_k) + c_{2i} E(X_2 X_k) + \dots + c_{mi} E(X_m X_k)$$

mas $E(X_1 X_k) = \sigma_{1k}$; $E(X_2 X_k) = \sigma_{2k}$, ... $E(X_m X_k) = \sigma_{mk}$

então:

$$\text{Cov}(X_k, C_i) = c_{1i} \sigma_{1k} + c_{2i} \sigma_{2k} + \dots + c_{mi} \sigma_{mk}$$

$$\text{cov}(X_k, C_i) = \sum_{k=1}^m c_{ki} \sigma_{k'k}$$

Considerando-se o conjunto de todas as variáveis X_k , $k=1,2,\dots,m$, as covariâncias entre esse conjunto e a componente C_i é dada pelo vetor

$$\sum c_i$$

onde \sum é a matriz de covariância das variáveis X , $c_i = (c_{1i}, c_{2i}, \dots, c_{mi})'$ é o i -ésima auto vetor de \sum (ou os coeficientes da combinação linear que formam C_i). Sendo c_i auto vetor de \sum tem-se:

$$\sum c_i = \lambda_i c_i$$

A covariância entre a variável X_k e a componente C_i será:

$$\text{Cov}(X_k, C_i) = \lambda_i c_{ik}$$

A correlação entre X_k e C_i será:

$$\text{Corr}(X_k, C_i) = \frac{\text{cov}(X_k, C_i)}{\sqrt{\text{var}(X_k) \text{var}(C_i)}} = \frac{\lambda_i c_{ik}}{\sigma_k \sqrt{\lambda_i}} = \frac{\sqrt{\lambda_i} c_{ik}}{\sigma_k}$$

Quando as variáveis originais forem padronizadas (isto é $\mu_k=0$ e $\sigma_k^2=1$, $k=1,2,\dots,m$) a correlação entre cada variável X_k e a componente C_i será:

$$\text{Cov}(X_k, C_i) = \lambda_i c_{ik}$$

Essas medidas de covariância e correlação entre variáveis e componentes possibilitam a interpretação das componentes principais através das variáveis que possuem as maiores cova

riâncias com cada componente.

A Agregação das Linhas

A partir da matriz A ($n \times m$) considere suas linhas X_ℓ ($\ell = 1, \dots, n$) com n variáveis aleatórias sobre as quais são feitas m observações. Trata-se nesse caso, de uma nuvem de m pontos no espaço R^n . As n componentes extraídas serão:

$$F_i = \sum_{\ell=1}^n f_{\ell i} X_\ell$$

Análogo ao caso da A.C.P. no espaço R^m , aqui trata-se de extrair n componentes não correlacionados e com variâncias decrescentes. Essas componentes são obtidas extraíndo-se os auto vetores e auto valores da matriz de covariância dos dados originais.

A Relação Entre as Duas Agregações

(Veja LEBART e FENELON - 1973)

Sendo X ($n \times m$) a matriz dos dados originais, a agregação das colunas é feita a partir da matriz de covariância $\sum_{\text{coluna}} = X'X$ ($m \times m$), e a agregação das linhas é feita a partir da matriz de covariância $\sum_{\text{linha}} = XX'$ ($n \times n$). Embora essas duas matrizes tenham dimensões diferentes elas tem o mesmo posto obviamente, e portanto o nº de auto valores não nulos será o mesmo para as duas matrizes. Além disso, os auto vetores e auto valores extraídos de $X'X$ e XX' são relacionados da seguinte forma:

Se c_i é auto vetor normalizado de $X'X$ com auto valor λ_i , pode-se escrever:

$$X'X c_i = \lambda_i c_i \quad \text{com } c_i' c_i = 1$$

Pré-multiplicando por X :

$$XX' (X c_i) = \lambda_i (X c_i)$$

Assim $X c_i$ é auto vetor de XX' com auto valor λ_i . Isto é, ao auto vetor e auto valor c_i e λ_i de $X'X$ corresponde um auto vetor $f_i^* = X c_i$ associado ao mesmo auto valor λ_i . A normalização de f_i^* é obtida por:

$$f_i^{*'} f_i^* = (X c_i)' (X c_i) = c_i' X' X c_i$$

mas por outro lado:

$$X'X c_i = \lambda_i c_i$$

$$c_i' X'X c_i = \lambda_i c_i' c_i = \lambda_i$$

portanto $f_i^{*'} f_i^* = \lambda_i$

Assim o auto vetor normalizado de XX' associado ao auto valor λ_i será:

$$f_i = \frac{1}{\sqrt{\lambda_i}} X c_i \quad (\text{II-3})$$

Usando o mesmo raciocínio, é possível se obter o auto vetor c_i de $X'X$ a partir de f_i e λ_i :

$$c_i = \frac{1}{\sqrt{\lambda_i}} X' f_i \quad (\text{II-4})$$

A Reconstrução dos Dados Iniciais X

A análise das componentes principais permite que se reconstrua a matriz de dados originais a partir das componentes que agregam linhas e colunas. Isto é, a partir do conhecimento de c_i , f_i e λ_i é possível reconstruir a matriz de dados inicial X , porque:

$$\sqrt{\lambda_i} f_i = X c_i$$

$$\sqrt{\lambda_i} f_i c_i' = X c_i c_i'$$

Somando para todo i

$$\sum_i \sqrt{\lambda_i} f_i c_i' = X \sum_i c_i c_i' = XI$$

Então:

$$X (n \times m) = \sum_{i=1}^m \sqrt{\lambda_i} f_i c_i' \quad (\text{supondo } m < n = \text{posto de } X)$$

No caso de se trabalhar com apenas as $p < m$ primeiras componentes a reconstrução é feita aproximadamente por:

$$X \approx \sum_{i=1}^p \sqrt{\lambda_i} f_i c_i'$$

Assim cada termo $a_{\ell k}$ ($\ell = 1, 2, \dots, n$; $k = 1, 2, \dots, m$) da matriz original será:

$$a_{\ell k} \approx \sum_{i=1}^p \sqrt{\lambda_i} f_{\ell i} c_{ik} \quad (\text{II-5})$$

Outras Abordagens da A.C.P.

Na descrição da técnica da A.C.P., foi usada a hipótese de que os dados da matriz X eram centrados, isto é as variáveis consideradas (X_k ou X_ℓ) têm média zero. Na verdade qualquer matriz de dados pode ser facilmente transformada para que satisfaça essa condição. Outras transformações dos dados são muitas vezes utilizadas dependendo da heterogeneidade das variáveis obtidas. Até aqui considerou-se que as componentes principais eram extraídas da matriz de covariância. Essas componentes são válidas quando se trabalha com uma matriz razoavelmente homogênea na qual as gradezas das variáveis são comparáveis de tal forma que a matriz de covariância é um instrumento adequado para a análise da estrutura dos dados. Nos casos em que a matriz de dados não é homogênea, apresentando variáveis medidas em diferentes unidades, é usual trabalhar-se com as variáveis padronizadas. Isto é, ao invés de tomar como ponto de partida a matriz X com

$$x_{k\ell} = X_{k\ell} - \bar{X}_k,$$

parte-se da matriz Z onde

$$z = \frac{X_{k\ell} - \bar{X}_k}{s_k}$$

\bar{X}_k é a média da k-ésima variável

s_k é o desvio padrão da k-ésima variável

Dessa forma a matriz $Z'Z$ a ser analisada é a matriz de correlação das variáveis. Em princípio pode parecer vantajoso trabalhar sempre com a matriz de correlação, que contorna o eventual problema da heterogeneidade dos dados, mas por outro lado, a inferência estatística sobre auto valores e auto vetores da matriz de correlação (criação de intervalos de confiança e testes de hipóteses) é muito mais complicada do que no caso em que se usa a matriz de covariância.

Outra transformação dos dados usada quando se trabalha com tabelas de contingência, é a proposta por BENZECRI (1973) na técnica chamada Análise de Correspondência. Os dados são transformados de tal modo que se parte da matriz R com

$$r_{ij} = \frac{p_{ij} - p_i \cdot p_j}{\sqrt{p_i p_j}}$$

onde

$$p_{ij} = \frac{X_{ij}}{\sum_{ij} X_{ij}}$$

$$p_i = \sum_j p_{ij}$$

$$p_j = \sum_i p_{ij}$$

Portanto se o objetivo é extrair-se a estrutura básica de uma matriz A , de acordo com os dados disponíveis aplica-se uma dessas transformações e procede-se a A.C.P. obtendo-se os auto valores e auto vetores de

$$X'X \text{ e } XX'$$

ou

$$Z'Z \text{ e } ZZ'$$

ou

$$R'R \text{ e } RR'$$

os quais formam o par de agregações ou a chamada estrutura básica da matriz.

É possível reduzir a série de matrizes A_1, A_2, \dots, A_T a série de estrutura básicas, e o problema de previsão passa a ser: obter a estrutura básica para o tempo $T + 1$ a partir da série de estruturas para $t = 1, 2, \dots, T$.

A matriz A_{T+1} é obtida aproximadamente pela fórmula de reconstrução da matriz de dados. Resta agora resolver o problema específico da previsão, já que, como foi dito anteriormente, as séries disponíveis são pequenas demais para a utilização dos métodos de previsão clássicos.

O problema será abordado através da idéia de estabilidade da série de matrizes. Isto é, uma vez que as séries são muito pequenas, a previsão só será possível se a evolução das

matrizes no tempo se der de forma bastante estável. O próximo capítulo é todo dedicado a criação de critérios de estabilidade, baseados nos coeficientes das matrizes ou baseados nas suas estruturas básicas. Em seguida, no capítulo IV, alguns procedimentos para previsão são propostos, além de avaliação dos erros cometidos, tanto na previsão quanto na reconstrução da matriz a partir de sua estrutura básica.

CAPÍTULO III

CRITÉRIO DE ESTABILIDADE

Nesse capítulo serão estabelecidos três tipos de critérios de estabilidade para a evolução da série de matrizes. Os primeiros critérios estabelecidos são baseados na evolução dos coeficientes das matrizes representados por sua média ou "centro de gravidade". Em seguida, a evolução da série de matrizes é observada através das suas estruturas básicas, isto é, as p primeiras componentes principais extraídas das matrizes A^t e suas transpostas. Essas estruturas serão primeiro analisadas diretamente, buscando-se verificar "semelhanças" entre elas nos diferentes tempos, e posteriormente a análise será feita através da agregação de todas as componentes C_i^t , $i = 1, 2, \dots, p$; $t = 1, 2, \dots, T$ (segundo nível de agregação) onde será possível estabelecer critérios de "semelhança" e "diferença" entre as componentes.

III.1 - Estabilidade do Centro de Gravidade

Como foi visto no Capítulo II, pode-se considerar a matriz econômica $A(n \times m)$ com elemento $a_{\ell k}$ ($\ell = 1, 2, \dots, n$; $k = 1, 2, \dots, m$), como uma matriz de dados com \underline{m} variáveis e \underline{n} observações. Cada uma das m colunas de A será uma variável X_k com observações $a_{\ell k}$ (k fixo, $\ell = 1, 2, \dots, n$). Nesse caso, é possível representar esses dados com uma nuvem de n pontos no espaço R^m .

Definição III.1: Centro de gravidade da nuvem de n pontos no R^m .

Considere a nuvem de n pontos no espaço R^m fornecida pela matriz A . Seja a média da variável X_k ($k = 1, 2 \dots m$) dada por:

$$\bar{X}_k = \frac{1}{n} \sum_{\ell=1}^n a_{\ell k}$$

chama-se centro de gravidade da nuvem de n pontos no espaço R^m ao ponto $CG \in R^m$ cujas coordenadas são as médias das m variáveis consideradas.

$$CG = (\bar{X}_1, \bar{X}_2, \dots, \bar{X}_m) .$$

Considerando-se a série de matrizes A^t ($t=1,2,\dots,T$), a cada uma delas corresponderá um centro de gravidade:

$$CG^t = (\bar{X}_1^t, \bar{X}_2^t, \dots, \bar{X}_m^t) ; \quad t = 1, 2, \dots, T$$

formando assim uma série de T pontos $CG^t \in R^m$.

Os primeiros critérios estabelecidos baseiam-se na evolução desses pontos CG^t . Inicialmente será analisado o modelo estacionário, isto é, o caso em que a série de pontos CG^t ($t = 1, 2, \dots, T$) sofre apenas pequenas alterações aleatórias de ano para ano. Este seria o caso em que não existe nenhum fator contribuindo para a mudança das médias das variáveis X_k ao longo do tempo. Portanto não haveria uma evolução propriamente dita, mas apenas perturbações aleatórias contribuindo para a variação dos centros de gravidade.

Modelo Estacionário para a Série dos Centros de Gravidade

Hipóteses básicas do modelo:

i) CG^t é uma variável aleatória $\in \mathbb{R}^m$ descrita por

ii) $CG^t = CG + \varepsilon^t$, onde:

iii) $CG \in \mathbb{R}^m$ é uma constante independente de t .

iv) $\varepsilon^t \in \mathbb{R}^m$ é a perturbação aleatória no ano t .

v) ε^t tem distribuição multinormal

vi) $E(\varepsilon^t) = 0 \quad \forall t$

vii) A matriz de covariância de ε^t , Σ_ε e constante isto é:

$$\text{var}(\varepsilon_k^t) = \sigma_k^2 \quad \forall t$$

$$\text{cov}(\varepsilon_k^t, \varepsilon_{k'}^t) = \sigma_{kk'} \quad \forall t$$

Sob essas hipóteses, a variável aleatória CG^t terá as seguintes propriedades:

i) CG^t tem distribuição multinormal

$$\text{ii) } E[CG^t] = E[CG] + E[\varepsilon^t] = CG$$

$$\text{iii) } \text{var}(CG_k^t) = \text{var}(CG_k) + \text{var}(\varepsilon_k^t) = \sigma_k^2$$

$$\text{iv) } \text{cov}(CG_k^t, CG_{k'}^t) = \text{cov}(\varepsilon_k^t, \varepsilon_{k'}^t) = \sigma_{kk'}$$

Assim a matriz de covariância de CG^t será

$$\sum CG^t = \sum_{\varepsilon} \text{constante } \forall t.$$

As estimativas para a média e para a matriz de co variância são dadas por:

$$\overline{CG} = \frac{1}{T} \sum_{t=1}^T CG^t$$

$$S_{CG}^2 = \frac{1}{T-1} \left((CG^t - \overline{CG}) (CG^t - \overline{CG})' \right)$$

$$\overline{CG}, CG^t \in R^m.$$

Embora possa parecer pouco provável que ocorra uma evolução estacionária para A^t , é importante chamar atenção para o fato que sendo A^t a matriz de insumo-produto, é razoável supor que não haja mudança nos $a_{\ell k}$ se não houver mudanças tecnológicas no período, mesmo se o nível de produção da economia apresentar uma real evolução no tempo. Isto porque nesse caso $a_{\ell k}$ representa o gasto por unidade de produto. Sempre que a matriz A^t for composta de elementos $a_{\ell k}$ "normalizados" por natureza, a hipótese do modelo estacionário não implica necessariamente numa economia estacionária.

Ajuste do Modelo Estacionário

Verificar se o modelo estacionário descreve adequadamente a evolução dos Centros de Gravidade no tempo, é verificar se as médias das variáveis X_k permanecem constantes no tempo porque:

- O modelo estacionário se ajusta quando $E(CG^1) = E(CG^2) = \dots = E(CG^T)$ onde $CG^t = (\bar{X}_1^t, \bar{X}_2^t, \dots, \bar{X}_m^t)$.

Por outro lado, considere o vetor X das m variáveis:

$$X \in \mathbb{R}^m = (X_1, X_2, \dots, X_m)$$

A cada variável X_k corresponde n observações $a_{\ell k}$, $\ell = 1, 2, \dots, n$, sendo μ^t o vetor das médias das m variáveis no tempo t , isto é:

$$\mu^t = (\mu_1^t, \mu_2^t, \dots, \mu_m^t)'$$

Se $E(CG^t)$ é constante $\forall t$, $E(\bar{X}^t)$ é constante $\forall t$. Mas $E(\bar{X}^t) = E(X^t) = \mu_t$. Portanto verificar se $E(CG^t)$ é constante $\forall t$ é equivalente a verificar se μ^t é constante $\forall t$, ou ainda verificar se:

$$\mu^1 = \mu^2 = \dots = \mu^T, \quad \mu^t \in \mathbb{R}^m$$

Dessa forma, é possível verificar o ajuste ao modelo estacionário através de um teste de igualdade de médias. A análise de variância multivariada é o tipo de testes adequado para verificar ou não essa igualdade.

($i = 1, 2, \dots, 5$) mostrada na tabela 14 do anexo. Os erros de previsão para F_1 , F_2 e F_5 são avaliados por:

i) Componente F_1

$$\text{Erro max} = \frac{r_1}{\|O_1\|} = \frac{.057}{.999} = .057$$

ii) Componente F_2

$$\text{Erro max} = \frac{r_2}{\|O_2\|} = \frac{.104}{.998} = .104$$

iii) Componente F_5

$$\text{Erro max} = \frac{r_5}{\|O_5\|} = \frac{.288}{.919} = .313$$

A reconstrução da matriz A^{81} foi obtida a partir das previsões C^{81} e F^{81} das tabelas 13 e 14 do anexo, e ainda a previsão dos auto valores (variâncias das componentes) para o ano de 81. Pela tabela (V.1) pode-se observar que os auto valores das cinco componentes consideradas, sofreram muito pouca alteração ao longo do tempo, e a previsão para 1981 pode ser obtida pela média ao longo do tempo. A reconstrução da matriz \bar{A} é dada por:

$$\bar{A}^{81} \cong \sum_{i=1}^5 \sqrt{\lambda_i^{81}} F_i^{81} \left(C_i^{81} \right)'$$

A matriz assim obtida \bar{A} é na verdade a matriz norma-

