OS VALORES SINGULARES NA ANÁLISE DE NÃO UNICIDADE E INSTABILIDADE:

Vamos considerar agora o funcional ajustante $\overline{\mathbf{y}}$ dado por $\overline{\mathbf{y}} = \overline{\overline{\mathbf{A}}}\ \overline{\mathbf{p}}$. Queremos estimar um vetor de parâmetros $\overline{\mathbf{p}}'$ que produza um funcional ajustante avaliado nos N pontos de observações que esteja perto das observações medidas $\overline{\mathbf{y}}^{o}$ sob algum critério. Se a medida de proximidade $\overline{\mathbf{y}}$ e $\overline{\mathbf{y}}^{o}$ for a distância Euclideana entre 2 vetores, o problema se torna:

$$\min_{\overline{\mathbf{p}}} \left\| \overline{\mathbf{y}}^{\mathbf{o}} - \overline{\overline{\mathbf{A}}}\ \overline{\mathbf{p}} \right\|_2^2 = \min_{\mathbf{p}} Q$$

Introduzindo a decomposição em valores singulares de $\overline{\overline{\mathbf{A}}}$:

$$Q = \| \overline{\boldsymbol{\varepsilon}} \|_2^2 = \left\| \overline{\mathbf{y}}^{\mathbf{o}} - \overline{\overline{\mathbf{U}}}\ \overline{\overline{\mathbf{S}}}\ \overline{\overline{\mathbf{V}}}^T\ \overline{\mathbf{p}} \right\|_2^2 \tag{6}$$

Propriedade: Numa transformação ortogonal a normal Euclideana é preservada. Assim por exemplo dado uma matriz ortogonal $\overline{\overline{\mathbf{U}}}$ então $\| \overline{\mathbf{x}} \|_2 = \| \overline{\overline{\mathbf{U}}}\ \overline{\mathbf{x}} \|_2$.

Vamos então aplicar esta propriedade na equação (6). Assim pré-multiplicando o vetor de resíduos pela matriz ortogonal $\overline{\overline{\mathbf{U}}}^T$ teremos a preservação da norma Euclideana do vetor de resíduos.

$$Q = \left\| \overline{\overline{\mathbf{U}}}^T\ \overline{\boldsymbol{\varepsilon}} \right\|_2^2 = \left\| \overline{\overline{\mathbf{U}}}^T\ \overline{\mathbf{y}}^{\mathbf{o}} - \overline{\overline{\mathbf{U}}}^T \overline{\overline{\mathbf{U}}}\ \overline{\overline{\mathbf{S}}}\ \overline{\overline{\mathbf{V}}}^T\ \overline{\mathbf{p}} \right\|_2^2 \tag{7}$$

resultando em :

$$Q = \left\| \overline{\overline{\mathbf{U}}}^T\ \overline{\mathbf{y}}^{\mathbf{o}} - \overline{\overline{\mathbf{S}}}\ \ \overline{\overline{\mathbf{V}}}^T\ \overline{\mathbf{p}} \right\|_2^2 \tag{8}$$

Vamos chamar:

$$\overline{\overline{\mathbf{U}}}^T \, \overline{\mathbf{y}}^o = \overline{\boldsymbol{\beta}} \tag{9}$$

e

$$\overline{\overline{\mathbf{V}}}^T \, \overline{\mathbf{p}} = \overline{\boldsymbol{\alpha}} \tag{10}$$

Substituindo as equações (9) e (10) na equação (8) obtemos que o nosso funcional Q a ser minimizado é expresso como:

$$Q = \left\| \, \overline{\boldsymbol{\beta}} - \overline{\overline{\mathbf{S}}} \, \overline{\boldsymbol{\alpha}} \, \right\|_2^2 \tag{11}$$

Matematicamente introduzimos um novo sistema de referência rotacionando os eixos originais para novos eixos de referência. Este novo sistema de referência tem como eixos os autovetores das matrizes ortogonais $\overline{\overline{\mathbf{U}}}$ e $\overline{\overline{\mathbf{V}}}$.

Considere que um vetor de parâmetros de dimensão 2, logo temos, p_1 e p_2. Então inicialmente temos o espaço de parâmetros formado pelos eixos p_1 e p_2. O ponto p na Figura (1) é o vetor $\overline{\mathbf{p}}$ que apresenta coordenadas ou elementos p_1 e p_2. Agora queremos representar o mesmo ponto p no novo sistema de referências, ou seja, em relação aos novos eixos formados pelos autovetores da matriz $\overline{\overline{\mathbf{V}}} = \left[\overline{\mathbf{v}}_1 \quad \overline{\mathbf{v}}_2 \right]$. Então o nosso objetivo é determinar as coordenadas do ponto p (elementos do vetor $\overline{\mathbf{p}}$) neste novo sistema de referência, ou seja, determinar α_1 e α_2 que são elementos do vetor $\overline{\boldsymbol{\alpha}}$. Geometricamente, temos:

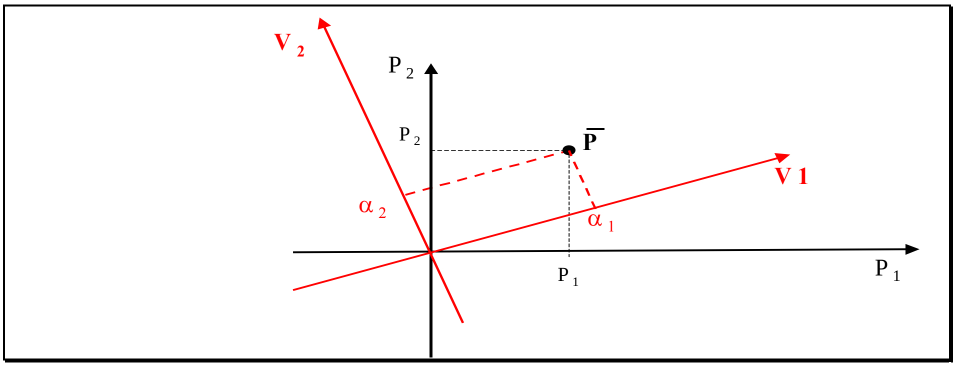


Figura 1

Matematicamente a representação das coordenadas antigas em relação aos eixos novos é expressa por

$$\overline{\overline{\mathbf{V}}}^T \ \overline{\mathbf{p}} = \overline{\boldsymbol{\alpha}} \qquad (12)$$

Considerando que o posto é igual a M então pré-multiplicando a equação acima pela matriz dos autovetores $\overline{\overline{\mathbf{V}}}$ temos que:

$$\overline{\mathbf{p}} = \overline{\overline{\mathbf{V}}} \ \overline{\boldsymbol{\alpha}} \qquad (13)$$

Analogamente se considerarmos que um vetor de observações de dimensão 2, logo temos, $y_1^{\,o}$ e $y_2^{\,o}$. Então inicialmente temos o espaço de observações formado pelos eixos $y_1^{\,o}$ e $y_2^{\,o}$. O ponto $y^{\,o}$ na Figura (2) é o vetor $\overline{\mathbf{y}}^{o}$ que apresenta coordenadas ou elementos $y_1^{\,o}$ e $y_2^{\,o}$. Agora queremos representar o mesmo ponto $y^{\,o}$ no novo sistema de referências, ou seja, em relação aos novos eixos formados pelos autovetores da matriz $\overline{\overline{U}} = \begin{bmatrix} \overline{u}_1 & \overline{u}_2 \end{bmatrix}$. Então o nosso objetivo é determinar as coordenadas do ponto $y^{\,o}$ (elementos do vetor $\overline{\mathbf{y}}^{o}$) neste novo sistema de referência, ou seja, determinar β_1 e β_2 que são elementos do vetor $\overline{\boldsymbol{\beta}}$. Geometricamente, temos:

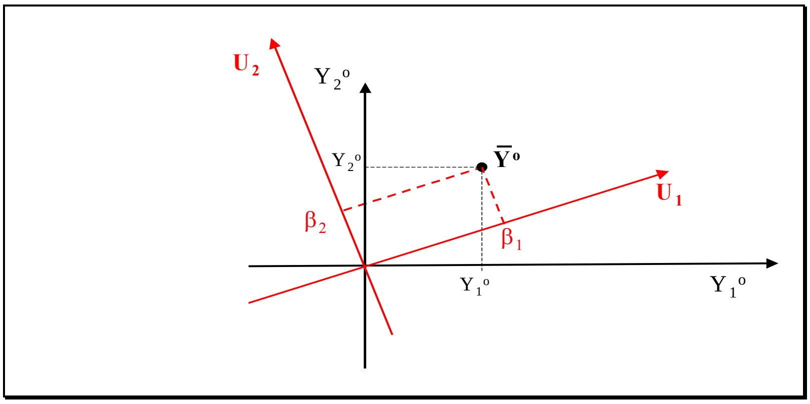


Figura 2

Matematicamente a representação das coordenadas antigas em relação aos eixos novos é expressa por

$$\overline{\overline{\mathbf{U}}}^T \, \overline{\mathbf{y}}^o = \overline{\boldsymbol{\beta}} \tag{14}$$

Considerando que o posto é igual a N então pré-multiplicando a equação acima pela matriz dos autovetores $\overline{\overline{\mathbf{U}}}$ temos que:

$$\overline{\mathbf{y}}^o = \overline{\overline{\mathbf{U}}} \, \overline{\boldsymbol{\beta}} \tag{15}$$

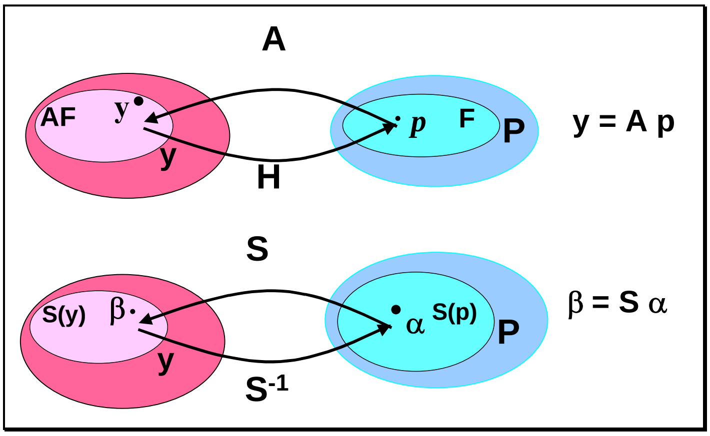


Figura 3

Dada uma matriz M x M $\overline{\overline{\mathbf{V}}} = \left[\overline{\mathbf{v}}_1 \quad \overline{\mathbf{v}}_2 \quad \dots \quad \overline{\mathbf{v}}_M \right]$ e um vetor M x 1 $\overline{\boldsymbol{\alpha}}$, temos que:

$$\overline{\overline{\mathbf{V}}} \, \overline{\boldsymbol{\alpha}} = \sum_{i=1}^{M} \alpha_i \overline{\mathbf{v}}_i \tag{16}$$

então

$$\overline{\overline{\mathbf{V}}} \, \overline{\boldsymbol{\alpha}} = \alpha_1 \overline{\mathbf{v}}_1 + \alpha_2 \overline{\mathbf{v}}_2 + \alpha_3 \overline{\mathbf{v}}_3 + \dots + \alpha_M \overline{\mathbf{v}}_M \tag{17}$$

Como os autovetores $\overline{\mathbf{v}}_i$, colunas da matiz $\overline{\overline{\mathbf{V}}}$ formam uma base no espaço de parâmetros, qualquer vetor de parâmetros $\overline{\mathbf{p}}'$ deste espaço que explica os dados geofísicos observados, ou seja satisfaz ao sistema $\overline{\mathbf{y}}^o = \overline{\overline{\mathbf{A}}} \, \overline{\mathbf{p}}'$, pode ser expandido como:

$$\bar{\mathbf{p}}' = \bar{\bar{\mathbf{V}}}\ \bar{\boldsymbol{\alpha}} = \sum_{i=1}^{M} \alpha_i \bar{\mathbf{v}}_i \tag{18}$$

Considerando agora a partição da matriz $\bar{\bar{\mathbf{V}}} = \left[\bar{\bar{\mathbf{V}}}_r \quad \bar{\bar{\mathbf{V}}}_{M-r}\right]$ temos:

$$\bar{\mathbf{p}}' = \left[\bar{\bar{\mathbf{V}}}_r \quad \bar{\bar{\mathbf{V}}}_{M-r}\right] \begin{bmatrix} \bar{\boldsymbol{\alpha}}_r \\ \bar{\boldsymbol{\alpha}}_{M-r} \end{bmatrix} \tag{19}$$

sendo que $\bar{\bar{\mathbf{V}}}_r$ **(M x r)** , $\bar{\bar{\mathbf{V}}}_{M-r}$ **(M x M - r)** , $\bar{\boldsymbol{\alpha}}_r$ **(r x 1)** e $\bar{\boldsymbol{\alpha}}_{M-r}$ **(M - r x 1)** .

Analogamente, os autovetores $\bar{\mathbf{u}}_i$ **,** $i = 1,\ldots N$, colunas da matiz $\bar{\bar{\mathbf{U}}}$ formam uma base no espaço das observações geofísicas. Então qualquer elemento deste espaço pode ser expandido como:

$$\bar{\mathbf{y}}^{\mathbf{o}} = \bar{\bar{\mathbf{U}}}\ \bar{\boldsymbol{\beta}} = \sum_{i=1}^{N} \beta_i \bar{\mathbf{u}}_i \tag{20}$$

Considerando N > M > r , logo o posto r é menor que M e N, e portanto teremos uma partição da matriz $\bar{\bar{\mathbf{U}}} = \left[\bar{\bar{\mathbf{U}}}_r \quad \bar{\bar{\mathbf{U}}}_{M-r}\ \bar{\bar{\mathbf{U}}}_{N-M}\right]$ e conseqüentemente temos

$$\bar{\mathbf{y}}^{\mathbf{o}} = \left[\bar{\bar{\mathbf{U}}}_r \quad \bar{\bar{\mathbf{U}}}_{M-r}\ \bar{\bar{\mathbf{U}}}_{N-M}\right] \begin{bmatrix} \bar{\boldsymbol{\beta}}_r \\ \bar{\boldsymbol{\beta}}_{M-r} \\ \bar{\boldsymbol{\beta}}_{N-M} \end{bmatrix} \tag{21}$$

sendo que

$\bar{\bar{\mathbf{U}}}_r$ **(N x r)**

$\bar{\bar{\mathbf{U}}}_{M-r}$ **(N x M - r)**

$\bar{\bar{\mathbf{U}}}_{N\text{-}M}$ **(Nx N-M)**

$\bar{\boldsymbol{\beta}}_r$ **(r x 1)**

$\bar{\boldsymbol{\beta}}_{M-r}$ **(M - r x 1)** .

$\bar{\boldsymbol{\beta}}_{N-M}$ **(N - M x 1)**

Vamos agora introduzir na equação (11) as partições da matriz $\bar{\bar{\mathbf{S}}}$ e dos vetores $\bar{\boldsymbol{\alpha}}$ e $\bar{\boldsymbol{\beta}}$ considerando N>M>r :

$$\mathbf{Q} = \left\| \bar{\boldsymbol{\beta}} - \bar{\bar{\mathbf{S}}}\ \bar{\boldsymbol{\alpha}} \right\|_2^2 = \left\| \begin{bmatrix} \bar{\boldsymbol{\beta}}_r \\ \bar{\boldsymbol{\beta}}_{M-r} \\ \bar{\boldsymbol{\beta}}_{N-M} \end{bmatrix} - \begin{bmatrix} \bar{\bar{\mathbf{S}}}_r & \bar{\bar{\mathbf{0}}} \\ \bar{\bar{\mathbf{0}}} & \bar{\bar{\mathbf{S}}}_{M-r} \\ \bar{\bar{\mathbf{0}}} & \bar{\bar{\mathbf{0}}} \end{bmatrix} \begin{bmatrix} \bar{\boldsymbol{\alpha}}_r \\ \bar{\boldsymbol{\alpha}}_{M-r} \end{bmatrix} \right\|_2^2 \qquad (22)$$

$$\mathbf{Q} = \left\| \begin{bmatrix} \bar{\boldsymbol{\beta}}_r \\ \bar{\boldsymbol{\beta}}_{M-r} \\ \bar{\boldsymbol{\beta}}_{N-M} \end{bmatrix} - \begin{bmatrix} \bar{\bar{\mathbf{S}}}_r\ \bar{\boldsymbol{\alpha}}_r \\ \bar{\bar{\mathbf{S}}}_{M-r}\ \bar{\boldsymbol{\alpha}}_{M-r} \\ \bar{\mathbf{0}} \end{bmatrix} \right\|_2^2$$

$$\mathbf{Q} = \left\| \begin{array}{c} \bar{\boldsymbol{\beta}}_r - \bar{\bar{\mathbf{S}}}_r \ \bar{\boldsymbol{\alpha}}_r \\ \bar{\boldsymbol{\beta}}_{M-r} - \bar{\bar{\mathbf{S}}}_{M-r} \ \bar{\boldsymbol{\alpha}}_{M-r} \\ \bar{\boldsymbol{\beta}}_{N-M} \end{array} \right\|_2^2$$

Finalmente obtemos:

$$Q = \left\| \bar{\boldsymbol{\beta}}_r - \bar{\bar{\mathbf{S}}}_r \ \bar{\boldsymbol{\alpha}}_r \right\|_2^2 + \left\| \bar{\boldsymbol{\beta}}_{M-r} - \bar{\bar{\mathbf{S}}}_{M-r} \ \bar{\boldsymbol{\alpha}}_{M-r} \right\|_2^2 + \left\| \bar{\boldsymbol{\beta}}_{N-M} \right\|_2^2 \quad (23)$$

A equação (23) é de extrema importância na análise da relação entre valores singulares, não unicidade e instabilidade.

Queremos encontrar uma solução $\bar{\mathbf{p}}'$ tal que a distância entre $\bar{\bar{\mathbf{A}}}\ \bar{\mathbf{p}}'$ e $\bar{\mathbf{y}}^{o}$ seja a menor possível na Norma Euclideana. Esta distância como vimos é dada pelo funcional Q. Para então obtermos a estimativa $\bar{\mathbf{p}}'$, basta, segundo a equação (18), determinarmos o vetor $\bar{\boldsymbol{\alpha}}$, uma vez que a matriz dos autovetores $\bar{\bar{\mathbf{V}}}$ é conhecida da decomposição em valores singulares da matriz $\bar{\bar{\mathbf{A}}}$.

Vejamos agora as possíveis hipóteses:

Hipótese 1: M – r valores singulares são nulos:

Neste caso a matriz $\bar{\bar{\mathbf{S}}}_{M-r} = \bar{\bar{\mathbf{0}}}$ então o funcional Q é dado por:

$$Q = \left\| \bar{\boldsymbol{\beta}}_r - \bar{\bar{\mathbf{S}}}_r \ \bar{\boldsymbol{\alpha}}_r \right\|_2^2 + \left\| \bar{\boldsymbol{\beta}}_{M-r} \right\|_2^2 + \left\| \bar{\boldsymbol{\beta}}_{N-M} \right\|_2^2 \qquad (24)$$

ou ainda

$$Q = \left(\bar{\boldsymbol{\beta}}_r - \bar{\bar{\mathbf{S}}}_r \ \bar{\boldsymbol{\alpha}}_r \right)^T \left(\bar{\boldsymbol{\beta}}_r - \bar{\bar{\mathbf{S}}}_r \ \bar{\boldsymbol{\alpha}}_r \right) + \left(\bar{\boldsymbol{\beta}}_{M-r} \right)^T \left(\bar{\boldsymbol{\beta}}_{M-r} \right) + \left(\bar{\boldsymbol{\beta}}_{N-M} \right)^T \left(\bar{\boldsymbol{\beta}}_{N-M} \right) \qquad (25)$$

Então queremos encontrar a estimativa $\bar{\boldsymbol{\alpha}}'$ tal que o funcional Q seja mínimo.

A condição necessária para Q ter um mínimo é

$$\overline{\nabla}_{\alpha}\ \mathbf{Q} = 0$$

A solução que minimiza Q nesta condição em que $\overline{\overline{\mathbf{S}}}_{M-r} = \overline{\overline{\mathbf{0}}}$ é dado por

$$\overline{\boldsymbol{\alpha}}_r{}' = \overline{\overline{\mathbf{S}}}_r{}^{-1}\ \overline{\boldsymbol{\beta}}_r \qquad (26)$$

Então a estimativa $\overline{\boldsymbol{\alpha}}'$ tal que o funcional Q seja mínimo é dada pela equação (26). Note que esta estimativa envolve apenas a parte do vetor $\overline{\boldsymbol{\alpha}}$ correspondente a $\overline{\boldsymbol{\alpha}}_r$, portanto a parte do vetor $\overline{\boldsymbol{\alpha}}$ correspondente a $\overline{\boldsymbol{\alpha}}_{M-r}$ não é usada para minimizar o funcional Q. Vimos na equação (18) que

$$\overline{\mathbf{p}}' = \overline{\overline{\mathbf{V}}}\ \overline{\boldsymbol{\alpha}}' = \overline{\overline{\mathbf{V}}}_r\ \overline{\boldsymbol{\alpha}}'_r + \overline{\overline{\mathbf{V}}}_{M-r}\ \overline{\boldsymbol{\alpha}}'_{M-r}\ .$$

Desta forma, qualquer vetor $\overline{\boldsymbol{\alpha}}'_{M-r}$ quando usado junto com o vetor $\overline{\boldsymbol{\alpha}}'_r$ dado pela equação (26) levará a vetores soluções:

$$\overline{\mathbf{p}}' = \overline{\overline{\mathbf{V}}}_r\ \overline{\boldsymbol{\alpha}}'_r \qquad + \qquad \overline{\overline{\mathbf{V}}}_{M-r}\ \overline{\boldsymbol{\alpha}}'_{M-r}$$

que produzirão mínimos da função objeto Q.

Substituindo a estimativa do vetor $\overline{\boldsymbol{\alpha}}$ dada pela equação (26) na equação (23) temos:

$$\mathbf{Q}_{\mathbf{MIN}} = \left\|\overline{\boldsymbol{\beta}}_{M-r}\right\|_2^2 + \left\|\ \overline{\boldsymbol{\beta}}_{N-M}\ \right\|_2^2 = \left\|\ \overline{\boldsymbol{\beta}}_{N-r}\ \right\|_2^2 \qquad (27)$$

Como, neste caso de M - r valores singulares nulos, podemos atribuir ao vetor $\overline{\boldsymbol{\alpha}}'_{M-r}$ qualquer valor, então teremos por conseqüência infinitas soluções $\overline{\mathbf{p}}' = \overline{\overline{\mathbf{V}}}_r\ \overline{\overline{\mathbf{S}}}_r{}^{-1}\ \overline{\boldsymbol{\beta}}_r + \overline{\overline{\mathbf{V}}}_{M-r}\ \overline{\boldsymbol{\alpha}}'_{M-r}$ produzindo o mesmo mínimo da função Q dado pela equação (27).

Note que o vetor $\bar{\boldsymbol{\alpha}}'_{M-r}$ **é a projeção da solução nula na base do espaço nulo do espaço dos parâmetros.** Então podemos dizer que

$$\bar{\mathbf{p}}^{null} = \bar{\bar{\mathbf{V}}}_{M-r}\ \bar{\boldsymbol{\alpha}}'_{M-r}$$

isto porque como consideramos que a matriz $\bar{\bar{\mathbf{A}}}$ tem M-r valores singulares iguais a ZERO, então sua SVD é dada por

$$\bar{\bar{\mathbf{A}}} = \begin{bmatrix} \bar{\bar{\mathbf{U}}}_r & \bar{\bar{\mathbf{U}}}_{M-r} & \bar{\bar{\mathbf{U}}}_{N-M} \end{bmatrix} \begin{bmatrix} \bar{\bar{\mathbf{S}}}_r & \bar{\bar{\mathbf{0}}} \\ \bar{\bar{\mathbf{0}}} & \bar{\bar{\mathbf{S}}}_{M-r} \\ \bar{\bar{\mathbf{0}}} & \bar{\bar{\mathbf{0}}} \end{bmatrix} \begin{bmatrix} \bar{\bar{\mathbf{V}}}_r^T \\ \bar{\bar{\mathbf{V}}}_{M-r}^T \end{bmatrix}$$

Sob a condição de $\bar{\bar{\mathbf{S}}}_{M-r} = \bar{\bar{\mathbf{0}}}$ pode-se escrever que

$$\bar{\bar{\mathbf{A}}} = \bar{\bar{\mathbf{U}}}_r\,\bar{\bar{\mathbf{S}}}_r\,\bar{\bar{\mathbf{V}}}_r^T$$

Então fazendo-se a multiplicação da matriz $\bar{\bar{\mathbf{A}}}$ por $\bar{\mathbf{p}}^{null} = \bar{\bar{\mathbf{V}}}_{M-r}\ \bar{\boldsymbol{\alpha}}'_{M-r}$ temos

$$\bar{\bar{\mathbf{A}}}\,\bar{\mathbf{p}}^{null} = \bar{\bar{\mathbf{U}}}_r\,\bar{\bar{\mathbf{S}}}_r\,\bar{\bar{\mathbf{V}}}_r^T\,\bar{\bar{\mathbf{V}}}_{M-r}\ \bar{\boldsymbol{\alpha}}'_{M-r}$$

Como $\bar{\bar{\mathbf{V}}}_r$ e $\bar{\bar{\mathbf{V}}}_{M-r}$ são matrizes ortogonais então $\bar{\bar{\mathbf{V}}}_r^T\,\bar{\bar{\mathbf{V}}}_{M-r} = \bar{\bar{\mathbf{0}}}_{(r\times r)}$

Conseqüentemente temos que

$$\bar{\bar{\mathbf{A}}}\,\bar{\mathbf{p}}^{null} = \bar{\mathbf{0}}$$

Provando que $\bar{\bar{\mathbf{V}}}_{M-r}\ \bar{\boldsymbol{\alpha}}'_{M-r}$ é um vetor solução nula

$$\bar{\bar{\mathbf{A}}}\,\bar{\mathbf{p}}^{null} = \bar{\bar{\mathbf{U}}}_r\,\bar{\bar{\mathbf{S}}}_r\,\bar{\bar{\mathbf{V}}}_r^T\,\bar{\bar{\mathbf{V}}}_{M-r}\ \bar{\boldsymbol{\alpha}}'_{M-r} = \bar{\mathbf{0}}$$

Então sob a hipótese de **M – r valores singulares nulos** temos a existência do espaço nulo e conseqüentemente M-r vetores soluções nulas, implicando uma solução geral dada por

$$\bar{\mathbf{p}}' = \bar{\bar{\mathbf{V}}}_r\ \bar{\boldsymbol{\alpha}}'_r + \bar{\bar{\mathbf{V}}}_{M-r}\ \bar{\boldsymbol{\alpha}}'_{M-r}$$

ou ainda

$$\overline{\mathbf{p}}_{geral} = \overline{\mathbf{p}}_{particular} + \overline{\mathbf{p}}^{null}$$

CONCLUSÃO:

Sob a hipótese de **M – r valores singulares nulos** temos a **não unicidade de soluções** produzindo o **mesmo ajuste** aos dados.

Hipótese 2: M – r valores singulares muito próximos de zero:

Então queremos encontrar a estimativa $\overline{\boldsymbol{\alpha}}'$ tal que o funcional Q seja mínimo, para isto vamos tomar $\overline{\nabla}_{\alpha}\, \mathbf{Q} = 0$. A solução que minimiza Q nesta condição em que temos M – r valores singulares próximos a zero é:

$$\overline{\boldsymbol{\alpha}}' = \overline{\overline{\mathbf{S}}}_M^{\mathbf{-1}}\, \overline{\boldsymbol{\beta}}_M = \begin{bmatrix} \overline{\boldsymbol{\alpha}}'_r \\ \overline{\boldsymbol{\alpha}}'_{M-r} \end{bmatrix} = \begin{bmatrix} \overline{\overline{\mathbf{S}}}_r^{\mathbf{-1}}\, \overline{\boldsymbol{\beta}}_r \\ \overline{\overline{\mathbf{S}}}_{M-r}^{\mathbf{-1}}\, \overline{\boldsymbol{\beta}}_{M-r} \end{bmatrix} \tag{28}$$

Vamos considerar que as observações geofísicas $\overline{\mathbf{y}}^o$ são exatas, ou seja sem contaminação de ruído. O vetor $\overline{\boldsymbol{\beta}}$ é dado por

$$\overline{\overline{\mathbf{U}}}^T\, \overline{\mathbf{y}}^o = \overline{\boldsymbol{\beta}} \tag{29}$$

Vimos na equação (18) que o vetor de parâmetros estimado será obtido resolvendo-se:

$$\overline{\mathbf{p}}' = \overline{\overline{\mathbf{V}}}\, \overline{\boldsymbol{\alpha}}' = \overline{\overline{\mathbf{V}}}_M\, \overline{\boldsymbol{\alpha}}'_M = \overline{\overline{\mathbf{V}}}_M\, \overline{\overline{\mathbf{S}}}_M^{-1}\, \overline{\boldsymbol{\beta}}_M \tag{30}$$

O vetor $\overline{\boldsymbol{\beta}}$ é exato e esta solução estimada é única

Vamos considerar que o nosso vetor de observações geofísicas $\overline{\mathbf{y}}^o$ estão contaminados por ruído:

$$\overline{\mathbf{y}} = \overline{\mathbf{y}}^o + \overline{\boldsymbol{\varepsilon}} \tag{31}$$

O vetor $\overline{\boldsymbol{\beta}}^n$ será :

$$\overline{\boldsymbol{\beta}}^n = \overline{\overline{\mathbf{U}}}^T \overline{\mathbf{y}}^n = \overline{\overline{\mathbf{U}}}^T \overline{\mathbf{y}}^o + \overline{\overline{\mathbf{U}}}^T \overline{\boldsymbol{\varepsilon}} \tag{32}$$

$$\overline{\boldsymbol{\beta}}^n = \overline{\boldsymbol{\beta}}_M + \overline{\boldsymbol{\beta}}_M^{Ruido} \tag{33}$$

sendo $\overline{\boldsymbol{\beta}}_M^{Ruido}$ é o vetor de valores aleatórios.

Vimos na equação (18) que o vetor de parâmetros estimado será obtido resolvendo-se:

$$\overline{\mathbf{p}}'' = \overline{\overline{\mathbf{V}}}\ \overline{\boldsymbol{\alpha}}' = \overline{\overline{\mathbf{V}}}_M\ \overline{\boldsymbol{\alpha}}'_M = \overline{\overline{\mathbf{V}}}_M\ \overline{\overline{\mathbf{S}}}_M^{-1} \overline{\boldsymbol{\beta}}^n \tag{34}$$

Substituindo-se a equação (33) na equação (34) temos:

$$\overline{\mathbf{p}}'' = \overline{\overline{\mathbf{V}}}_M\ \overline{\overline{\mathbf{S}}}_M^{-1} \overline{\boldsymbol{\beta}}_M + \overline{\overline{\mathbf{V}}}_M\ \overline{\overline{\mathbf{S}}}_M^{-1} \overline{\boldsymbol{\beta}}_M^{Ruido}$$

Como $\overline{\mathbf{p}}' = \overline{\overline{\mathbf{V}}}_M\ \overline{\overline{\mathbf{S}}}_M^{-1} \overline{\boldsymbol{\beta}}_M$ é a solução estimada para observações exatas (sem ruído) então, neste caso em que consideramos a existência de ruído temos:

$$\overline{\mathbf{p}}'' = \overline{\mathbf{p}}' + \overline{\overline{\mathbf{V}}}_M\ \overline{\overline{\mathbf{S}}}_M^{-1} \overline{\boldsymbol{\beta}}_M^{Ruido} \tag{35}$$

que pode ser rescrita como

$$\overline{\mathbf{p}}'' = \overline{\mathbf{p}}' + \sum_{i=1}^{M} \left(\frac{\beta_i^{Ruido}}{S_i} \right) \overline{\mathbf{v}}_{\mathbf{i}} \tag{36}$$

Note que a distância entre $\overline{\mathbf{p}}'$ e $\overline{\mathbf{p}}''$ será:

$$\left\|\overline{\mathbf{p}}' - \overline{\mathbf{p}}''\right\|^2 = \left\|\overline{\overline{\mathbf{V}}}_M \ \overline{\overline{\mathbf{S}}}_M^{-1} \overline{\boldsymbol{\beta}}_M^{Ruido}\right\|^2 \tag{37}$$

$$\left\|\overline{\mathbf{p}}' - \overline{\mathbf{p}}''\right\|^2 = \overline{\boldsymbol{\beta}}_M^{Ruido\,\mathrm{T}} \ \overline{\overline{\mathbf{S}}}_M^{-1} \overline{\overline{\mathbf{V}}}_M^T \overline{\overline{\mathbf{V}}}_M \ \overline{\overline{\mathbf{S}}}_M^{-1} \ \overline{\boldsymbol{\beta}}_M^{Ruido} \tag{38}$$

$$\left\|\overline{\mathbf{p}}' - \overline{\mathbf{p}}''\right\|^2 = \overline{\boldsymbol{\beta}}_M^{Ruido\ \mathrm{T}} \ \overline{\overline{\mathbf{S}}}_M^{-2} \ \overline{\boldsymbol{\beta}}_M^{Ruido} = \sum_{i=1}^{M} \left(\frac{\beta_i^{Ruido} \ \delta_i}{S_i} \right)^2 \tag{39}$$

$$\left\|\overline{\mathbf{p}}' - \overline{\mathbf{p}}''\right\|^2 = \sum_{i=1}^{r} \left(\frac{\beta_i^{Ruido}}{S_i} \right)^2 + \sum_{i=r+1}^{M} \left(\frac{\beta_i^{Ruido}}{S_i} \right)^2 \tag{40}$$

A equação (40) mostra que uma solução $\overline{\mathbf{p}}''$ estimada com dados contaminados por ruído e **sob a hipótese de M – r valores próximos a zero** pode estar infinitamente distante da solução estimada com dados exatos $\overline{\mathbf{p}}'$. Isto ocorre devido a segunda parcela do somatório da equação (40) $\left[\sum_{i=r+1}^{M} \left(\frac{\beta_i^{Ruido}}{S_i} \right)^2 \right]$; note que cada termo desta parcela pode apresentar valores grandes pois temos $S_i \to 0$ para $i = r+1, r+2, \ldots, M$.

Fazendo uma análise da equação (36) vemos que a solução estimada $\overline{\mathbf{p}}''$ é igual a solução estimada com dados exatos $\overline{\mathbf{p}}'$ somada a um vetor que envolve o ruído aleatório, os valores singulares e autovetores. Como β_i^{Ruido} é o ruído aleatório então pode assumir valores positivos ou negativos, os valores singulares $S_i, i = r+1, \ldots, M$ amplificarão este ruído aleatório fazendo com que as componentes da solução possam assumir valores bem elevados ora positivos ora negativos caracterizando assim a INSTABILIDADE DAS SOLUÇÕES.

CONCLUSÃO:

Sob a hipótese de **M – r valores singulares muito próximos a zero** temos a **instabilidade da solução** produzindo o **mesmo ajuste** aos dados.

OS VALORES SINGULARES NA ANÁLISE DO AJUSTE E DOS PARÂMETROS DETERMINADOS PELAS OBSERVAÇÕES:

Da análise dos valores singulares é possível detectar se um problema inverso tem solução não única ou se a solução é instável. Adicionalmente, esta análise permite:

1) Analisar se a solução estimada produzirá um ajuste exato dos dados (funcional ajustante passando exatamente pelas observações)
2) Descobrir quais os parâmetros ou combinações podem e quais não podem ser determinados pelas observações
3) Formular uma reparametrização do problema com vistas a estimar apenas os parâmetros (ou combinações de parâmetros) que podem ser determinados pelas observações

Ajuste exato:

Para a solução ser exata temos que obedecer a condição do resíduo Q ser igual a zero:

$$Q = 0$$

Se há M – r valores singulares nulos, a equação (27) fornece

$$Q_{\text{MIN}} = \left\| \bar{\boldsymbol{\beta}}_{M-r} \right\|_2^2 + \left\| \bar{\boldsymbol{\beta}}_{N-M} \right\|_2^2 = \left\| \bar{\boldsymbol{\beta}}_{N-r} \right\|_2^2$$

Note que o resíduo Q é igual ao norma Euclideana do vetor $\bar{\boldsymbol{\beta}}_{N-r}$. Como

$$\bar{\boldsymbol{\beta}}_{N-r} = \bar{\bar{\mathbf{U}}}_{N-r}^T \bar{\mathbf{y}}^o$$

a condição para a solução ser exata (Q=0) será:

$$\| \bar{\boldsymbol{\beta}}_{N-r} \|_2^2 = \| \bar{\bar{\mathbf{U}}}_{N-r}^T \bar{\mathbf{y}}^o \|_2^2 = 0$$

Isto ocorre em duas situações:

(1) r = N ≤ M

Neste caso a matriz $\bar{\bar{\mathbf{U}}}_{N-r} = \bar{\bar{\mathbf{0}}}$.

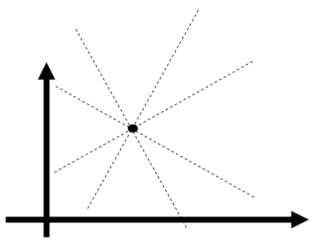
Vamos considerar o caso em que r= N=1, M=2 . Um exemplo poderia ser o ajuste de uma reta por um ponto yo. Neste exemplo, o dado observado é o único ponto yo e o nosso problema consiste em estimar os coeficientes angular e linear da reta (M=2) que explica este dado observado (yo). Temos, portanto, infinitas soluções e todas passando exatamente pelo ponto yo (ajuste exato).

Exemplo de um ajuste exato

Hipótese 1: M – r valores singulares são nulos $\bar{\bar{\mathbf{S}}}_{M-r} = \bar{\bar{\mathbf{0}}}$

Condição 1: se $\mathbf{r} = \mathbf{N} \leq \mathbf{M}$

N = 1 e M =2



(2) r < N

Vimos que a condição para a solução ser exata (Q=0) será:

$$\left\| \bar{\boldsymbol{\beta}}_{N-r} \right\|_2^2 = \left\| \bar{\bar{\mathbf{U}}}_{N-r}^T \bar{\mathbf{y}}^o \right\|_2^2 = 0$$

Como r < N isto ocorre se os autovetores que compõe a matriz $\bar{\bar{\mathbf{U}}}_{N-r}$ forem ortogonais ao vetor das observações $\bar{\mathbf{y}}^o$. Há, portanto, um total de N - r observações redundantes. Um exemplo poderia ser o ajuste de uma reta dado 5 pontos sendo que 3 pontos são colineares a 2 pontos quaisquer que são suficientes para ajustar uma reta. Geometricamente, isto corresponde ajustar uma reta dado N –3 observações colineares então apenas 2 observações são suficientes para ajustar uma reta. Note que no exemplo abaixo as observações 2 e 5, por exemplo, são suficientes para determinar a reta. Assim as observações 1, 3 e 4 são redundantes. Neste caso N=5 e r=2

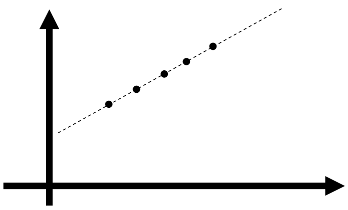
$$\begin{cases} \bar{\mathbf{u}}_3^T \bar{\mathbf{y}}^o = \bar{0} \\ \bar{\mathbf{u}}_4^T \bar{\mathbf{y}}^o = \bar{0} \\ \bar{\mathbf{u}}_5^T \bar{\mathbf{y}}^o = \bar{0} \end{cases}$$

Exemplo de um ajuste exato

Hipótese 1: M – r valores singulares são nulos $\bar{\bar{\mathbf{S}}}_{M-r} = \bar{\bar{\mathbf{0}}}$

Condição 2: **se r $<$ N**

Ajuste de uma reta M = 2, com N - 2 observações colineares



Quais os parâmetros ou combinações de parâmetros podem ser determinados:

Considere a equação (23):

$$\mathbf{Q} = \left\| \;\bar{\boldsymbol{\beta}}_r - \bar{\bar{\mathbf{S}}}_r\;\bar{\boldsymbol{\alpha}}_r \;\right\|_2^2 + \left\| \bar{\boldsymbol{\beta}}_{M-r} - \bar{\bar{\mathbf{S}}}_{M-r}\;\bar{\boldsymbol{\alpha}}_{M-r} \right\|_2^2 + \left\| \;\bar{\boldsymbol{\beta}}_{N-M}\; \right\|_2^2$$

Se existem M – r valores singulares nulos (ou muito próximos a zero) a matriz $\bar{\bar{\mathbf{S}}}_{M-r}$ será nula (ou virtualmente nula) de modo que o produto $\bar{\bar{\mathbf{S}}}_{M-r}\bar{\boldsymbol{\alpha}}_{M-r}$ será praticamente nulo. Assim a condição para a minimização de Q permite a penas determinar o vetor $\bar{\boldsymbol{\alpha}}_r$ pela relação:

$$\bar{\boldsymbol{\alpha}}_r{}' = \bar{\bar{\mathbf{S}}}_r^{\,\mathbf{-1}}\;\bar{\boldsymbol{\beta}}_r$$

Note que o vetor $\bar{\boldsymbol{\alpha}}_{M-r}$ não pode ser determinado da condição de minimização de Q, ou seja, $\bar{\boldsymbol{\alpha}}_{M-r}$ não pode ser estimado usando-se apenas as observações. Assim temos que

$$\begin{cases} \quad\bar{\boldsymbol{\alpha}}_r \text{ pode ser determinado} \\ \bar{\boldsymbol{\alpha}}_{M-r} \text{ não pode ser determinado} \end{cases}$$

Como $\bar{\boldsymbol{\alpha}}_r = \bar{\bar{\mathbf{V}}}_r^{\,T}\,\bar{\mathbf{p}}$ **e** $\bar{\boldsymbol{\alpha}}_{M-r} = \bar{\bar{\mathbf{V}}}_{M-r}^{T}\,\bar{\mathbf{p}}$, as combinações de parâmetros que podem ser determinadas são:

$$\bar{\mathbf{v}}_1^T\;\bar{\mathbf{p}},\quad \bar{\mathbf{v}}_2^T\;\bar{\mathbf{p}},\quad \dots,\; \bar{\mathbf{v}}_r^T\;\bar{\mathbf{p}} \quad .$$

As combinações de parâmetros que não podem ser determinados são:

$$\bar{\mathbf{v}}_{r+1}^T\;\bar{\mathbf{p}},\quad \bar{\mathbf{v}}_{r+2}^T\;\bar{\mathbf{p}},\quad \dots,\; \bar{\mathbf{v}}_M^T\;\bar{\mathbf{p}} \quad .$$

Veja então que podemos determinar dos dados observados geofísicos apenas a projeção dos parâmetros nas bases do espaço iluminado ($\bar{\bar{\mathbf{V}}}_r$ Figura 1) , ou seja, o vetor $\overline{\alpha}_r$ e NÃO podemos determinar dos dados observados geofísicos a projeção dos parâmetros nas bases do espaço não-iluminado (espaço nulo $\overline{\overline{\mathbf{V}}}_{M-r}$, Figura 1) .

Vamos retornar ao nosso exemplo simplificado que temos dois parâmetros e uma única observação relacionados pelo sistema de equações linear

$$\overline{\overline{\mathbf{A}}}\,\overline{\mathbf{p}} = \overline{\mathbf{y}}^{\mathbf{o}}$$

$$\begin{bmatrix} \frac{1}{2} & \frac{1}{2} \end{bmatrix} \begin{bmatrix} p_1 \\ p_2 \end{bmatrix} = y_1^o$$

Fazendo a SVD da matriz $\overline{\overline{\mathbf{A}}}$ temos que

$$\overline{\overline{\mathbf{U}}}_{(1\times1)} = \begin{bmatrix} \overline{\overline{\mathbf{U}}}_r \end{bmatrix} = 1$$

$$\overline{\overline{\mathbf{V}}}_{(2\times2)} = \begin{bmatrix} \frac{\sqrt{2}}{2} & -\frac{\sqrt{2}}{2} \\ \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \end{bmatrix} = \begin{bmatrix} \overline{\overline{\mathbf{V}}}_r & \overline{\overline{\mathbf{V}}}_{M-r} \end{bmatrix} \quad \text{e}$$

$$\overline{\overline{\mathbf{S}}}_{(1\times2)} = \begin{bmatrix} \frac{\sqrt{2}}{2} & 0 \end{bmatrix} = \begin{bmatrix} \overline{\overline{\mathbf{S}}}_r & \overline{\overline{\mathbf{S}}}_{M-r} \end{bmatrix}$$

Temos então M=2, N=1 e r=1. Há portanto, apenas um único valor singular diferente de zero. Já vimos que vetores colunas de $\overline{\overline{\mathbf{V}}}_r$ formam as bases do espaço iluminado e os vetores colunas de $\overline{\overline{\mathbf{V}}}_{M-r}$ formam as bases do espaço não-iluminado da matriz $\overline{\overline{\mathbf{A}}}$, neste problema as matrizes $\overline{\overline{\mathbf{V}}}_r$ e $\overline{\overline{\mathbf{V}}}_{M-r}$ tem apenas um único vetor coluna que são respectivamente:

$$\overline{\overline{\mathbf{V}}}_{r\,(2\times1)} = \begin{bmatrix} \frac{\sqrt{2}}{2} \\ \frac{\sqrt{2}}{2} \end{bmatrix} \text{ e } \overline{\overline{\mathbf{V}}}_{M-r\,(2\times1)} = \begin{bmatrix} -\frac{\sqrt{2}}{2} \\ \frac{\sqrt{2}}{2} \end{bmatrix}$$

As combinações de parâmetros que podem ser determinadas apenas a partir dos dados geofísicos observados são

$$\underset{(1\times2)}{\overline{\overline{\mathbf{V}}}_r^{\,T}} \; \underset{(2\times1)}{\overline{\mathbf{p}}} = \underset{(1\times1)}{\overline{\boldsymbol{\alpha}}_r} = \frac{\sqrt{2}\,(p_1 + p_2)}{2}$$

Portanto as combinações de parâmetros que podem ser determinadas apenas a partir dos dados geofísicos observados são, a menos que uma constante multiplicativa, as médias dos parâmetros p_1 e p_2.

Já as combinações de parâmetros que NÃO podem ser determinadas apenas a partir dos dados geofísicos observados são, a menos que uma constante multiplicativa, o contraste entre os parâmetros p_1 e p_2.

$$\underset{(1\times2)}{\overline{\overline{\mathbf{V}}}_{M-r}^{\,T}} \; \underset{(2\times1)}{\overline{\mathbf{p}}} = \underset{(1\times1)}{\overline{\boldsymbol{\alpha}}_{M-r}} = \frac{\sqrt{2}\,(p_1 - p_2)}{2}$$

<u>Reparametrização:</u>

Veja que a análise anterior mostrou que uma reparametrização imediata é fazer os novos parâmetros iguais a $\alpha_1, \alpha_2, \alpha_3, \ldots, \alpha_r$ que podem ser determinados APENAS dos dados (geofísicos) observados. O problema com esta medida é que com a reparametrização os novos parâmetros são, em geral, desprovido de significado físico no sistema de parâmetros originais. Lembre-se que

$$\underset{(r\,\times\,M)}{\overline{\overline{\mathbf{V}}}_r^{\,T}} \quad \underset{(M\,\times\,1)}{\overline{\mathbf{p}}} \quad = \quad \underset{(r\times 1)}{\overline{\boldsymbol{\alpha}}_r}$$

No entanto, os $\alpha_1, \alpha_2, \alpha_3, ..., \alpha_r$ podem guiar o intérprete para formular uma reparametrização que faça sentido físico

Um exemplo:

Considere que a matriz de sensibilidade é igual a:

$$\overline{\overline{\mathbf{A}}} = \begin{bmatrix} 1 & 1 \\ 1 & 1 \\ 1 & 1 \end{bmatrix}$$

A decomposição em valores singulares desta matriz é igual a:

$$\overline{\overline{\mathbf{A}}} = \overline{\overline{\mathbf{U}}}_r \; \overline{\overline{\mathbf{S}}}_r \; \overline{\overline{\mathbf{V}}}_r^{\,T} = \underbrace{\begin{pmatrix} \sqrt{3}/3 \\ \sqrt{3}/3 \\ \sqrt{3}/3 \end{pmatrix}}_{\overline{\overline{\mathbf{U}}}_r} \underbrace{\left(\sqrt{6}\right)}_{\overline{\overline{\mathbf{S}}}_r} \underbrace{\begin{pmatrix} \sqrt{2}/2 & \sqrt{2}/2 \end{pmatrix}}_{\overline{\overline{\mathbf{V}}}_r^{\,T}}$$

Já vimos que as combinações de parâmetros que podem ser determinadas apenas a partir dos dados geofísicos observados são dadas por

$$\underset{(1\,\times\,2)}{\overline{\overline{\mathbf{V}}}_r^{\,T}} \quad \underset{(2\,\times\,1)}{\overline{\mathbf{p}}} \quad = \quad \underset{(1\times 1)}{\overline{\boldsymbol{\alpha}}_r} \quad = \quad \frac{\sqrt{2}\,(p_1 + p_2)}{2}$$

Vamos achar agora achar as combinações de OBSERVAÇÕES que são LI

$$\overline{\overline{\mathbf{U}}}_r^{\,T}\, \overline{\mathbf{y}}^{\,o} = \overline{\boldsymbol{\beta}}_r$$

$$\overline{\overline{\mathbf{U}}}_r^{\,T}\,\overline{\mathbf{y}}^o = \overline{\boldsymbol{\beta}}_r = \begin{bmatrix} \sqrt{3}/3 & \sqrt{3}/3 & \sqrt{3}/3 \end{bmatrix} \begin{pmatrix} y_1^0 \\ y_2^0 \\ y_3^0 \end{pmatrix} = \frac{\sqrt{3}\left(y_1^0 + y_2^0 + y_3^0\right)}{3}$$

Para determinar a combinação de parâmetros $\left(\overline{\boldsymbol{\alpha}}_r\right)$ que podem ser determinadas apenas a partir dos dados geofísicos observados, não precisamos de todas as 3 observações, é necessário apenas as uma combinação de observações LI $\left(\overline{\boldsymbol{\beta}}_r\right)$.

Reparametrização:

O nosso sistema anterior é definido por

$$\overline{\overline{\mathbf{A}}}\;\overline{\mathbf{p}} = \overline{\mathbf{y}}^0$$

$$\begin{bmatrix} 1 & 1 \\ 1 & 1 \\ 1 & 1 \end{bmatrix} \begin{pmatrix} p_1 \\ p_2 \end{pmatrix} = \begin{pmatrix} y_1^0 \\ y_2^0 \\ y_3^0 \end{pmatrix}$$

Portanto eu tenho dois parâmetros para serem estimados. Vamos definir um NOVO vetor de parâmetros de dimensão (1 x 1) :

$$p_1' = \frac{\left(p_1 + p_2\right)}{2} \approx \alpha_r$$

Vamos definir um NOVO vetor de observações de dimensão (1 x 1):

$$y_1^{0\,\prime} = \frac{\left(y_1^0 + y_2^0 + y_3^0\right)}{3} \approx \beta_r$$

Um exemplo do uso da SVD como ferramenta de análise da não unicidade e instabilidade

Análise do estimador de Mínimos Quadrados Via decomposição em valores singulares da matriz de sensibilidade

A solução $\hat{\mathbf{p}}$ que chamamos de estimador de mínimos quadrados é

$$\hat{\mathbf{p}} = \left(\bar{\bar{\mathbf{A}}}^{\mathbf{T}}\bar{\bar{\mathbf{A}}}\right)^{-1}\bar{\bar{\mathbf{A}}}^{\mathbf{T}}\bar{\mathbf{y}}^{\mathbf{o}}$$

Consideraremos aqui que a matriz de sensibilidade $\bar{\bar{\mathbf{A}}}$ (N x M) é de posto completo[1] . Como o posto de uma matriz que é obtida pelo produto de duas outras matrizes é no máximo igual ao menor posto das matrizes, então temos que o posto de $\bar{\bar{\mathbf{A}}}^{\mathbf{T}}\bar{\bar{\mathbf{A}}} \leq \min(M,N)$ como presumimos que $\bar{\bar{\mathbf{A}}}$ tem posto completo, então o posto de $\bar{\bar{\mathbf{A}}}^{\mathbf{T}}\bar{\bar{\mathbf{A}}} = \min(M,N)$. Já vimos anteriormente que para usarmos o estimador de mínimos quadrados é necessário, que $N \geq M$, isto é que o número de observações (N) seja pelo menos igual ao número de parâmetros (M). Como neste caso consideramos que a matriz $\bar{\bar{\mathbf{A}}}$ é de posto completo temos que o posto da matriz $\bar{\bar{\mathbf{A}}}^{\mathbf{T}}\bar{\bar{\mathbf{A}}}$ é igual a M. Em outras palavras, se o $\mathbf{r}\left(\bar{\bar{\mathbf{A}}}^{\mathbf{T}}\bar{\bar{\mathbf{A}}}\right) = \mathbf{M}$ então há M valores singulares diferentes de zero. Então sob estas condições ($N \geq M = r$) a decomposição em valores singulares da matriz de sensibilidade é igual a $\bar{\bar{\mathbf{A}}} = \bar{\bar{\mathbf{U}}}_{\mathbf{r}}\ \bar{\bar{\mathbf{S}}}_{\mathbf{r}}\ \bar{\bar{\mathbf{V}}}_{\mathbf{r}}^{\mathbf{T}}$ em que r = M. Usando esta informação no estiamdor $\hat{\mathbf{p}}$

$$\hat{\mathbf{p}} = \left(\bar{\bar{\mathbf{A}}}^{\mathbf{T}}\bar{\bar{\mathbf{A}}}\right)^{-1}\bar{\bar{\mathbf{A}}}^{\mathbf{T}}\bar{\mathbf{y}}^{\mathbf{o}}$$

temos

$$\hat{\mathbf{p}} = \left(\bar{\bar{\mathbf{V}}}_{\mathrm{M}} \bar{\bar{\mathbf{S}}}_{\mathrm{M}} \bar{\bar{\mathbf{U}}}_{\mathrm{M}}^{\mathrm{T}} \bar{\bar{\mathbf{U}}}_{\mathrm{M}} \bar{\bar{\mathbf{S}}}_{\mathrm{M}} \bar{\bar{\mathbf{V}}}_{\mathrm{M}}^{\mathrm{T}} \right)^{-1} \bar{\bar{\mathbf{V}}}_{\mathrm{M}} \bar{\bar{\mathbf{S}}}_{\mathrm{M}} \bar{\bar{\mathbf{U}}}_{\mathrm{M}}^{\mathrm{T}} \bar{\mathbf{y}}^{o}$$

$$\hat{\mathbf{p}} = \left(\bar{\bar{\mathbf{V}}}_{\mathrm{M}} \bar{\bar{\mathbf{S}}}_{\mathrm{M}}^{2} \bar{\bar{\mathbf{V}}}_{\mathrm{M}}^{\mathrm{T}} \right)^{-1} \bar{\bar{\mathbf{V}}}_{\mathrm{M}} \bar{\bar{\mathbf{S}}}_{\mathrm{M}} \bar{\bar{\mathbf{U}}}_{\mathrm{M}}^{\mathrm{T}} \bar{\mathbf{y}}^{o}$$

como r = M temos que $\bar{\bar{\mathbf{V}}}_{M}^{-1} = \bar{\bar{\mathbf{V}}}_{M}^{\mathrm{T}}$ então temos que

$$\hat{\mathbf{p}} = \bar{\bar{\mathbf{V}}}_{\mathrm{M}} \bar{\bar{\mathbf{S}}}_{\mathrm{M}}^{-2} \bar{\bar{\mathbf{V}}}_{\mathrm{M}}^{\mathrm{T}} \bar{\bar{\mathbf{V}}}_{\mathrm{M}} \bar{\bar{\mathbf{S}}}_{\mathrm{M}} \bar{\bar{\mathbf{U}}}_{\mathrm{M}}^{\mathrm{T}} \bar{\mathbf{y}}^{o}$$

como $\bar{\bar{\mathbf{V}}}^{\mathbf{T}}_{r} \bar{\bar{\mathbf{V}}}_{r} = \bar{\bar{\mathbf{I}}}_{(r \times r)}$ e $\bar{\bar{\mathbf{V}}}^{\mathbf{T}}_{M-r} \bar{\bar{\mathbf{V}}}_{M-r} = \bar{\bar{\mathbf{I}}}_{(M-r \times M-r)}$ então temos que $\bar{\bar{\mathbf{V}}}_{M}^{T} \bar{\bar{\mathbf{V}}}_{M} = \bar{\bar{\mathbf{I}}}_{M}$

$$\hat{\mathbf{p}} = \bar{\bar{\mathbf{V}}}_{\mathrm{M}} \bar{\bar{\mathbf{S}}}_{\mathrm{M}}^{-2} \bar{\bar{\mathbf{S}}}_{\mathrm{M}} \bar{\bar{\mathbf{U}}}_{\mathrm{M}}^{\mathrm{T}} \bar{\mathbf{y}}^{o}$$

$$\bar{\hat{\mathbf{p}}} = \bar{\bar{\mathbf{V}}}_{\mathrm{M}} \bar{\bar{\mathbf{S}}}_{\mathrm{M}}^{-1} \bar{\bar{\mathbf{U}}}_{\mathrm{M}}^{\mathrm{T}} \bar{\mathbf{y}}^{o}$$

Vamos chamar: o vetor $\bar{\bar{\mathbf{U}}}_{M}^{T} \bar{\mathbf{y}}^{o} = \bar{\boldsymbol{\beta}}$ (M x 1) então

$$\bar{\hat{\mathbf{p}}} = \bar{\bar{\mathbf{V}}}_{\mathrm{M}} \bar{\bar{\mathbf{S}}}_{\mathrm{M}}^{-1} \bar{\boldsymbol{\beta}}$$

que pode ser escrito na forma de somatório como

$$\bar{\hat{\mathbf{p}}} = \sum_{\mathrm{i}=1}^{\mathrm{M}} \bar{\mathbf{v}}_{i} \frac{\beta_{i}}{S_{i}}$$

em que $\bar{\mathbf{v}}_{i}, i = 1,2,\ldots M$ são os M vetores colunas que formam a matriz dos autovetores $\bar{\bar{\mathbf{V}}}_{M} \, (M \times M)$. Veja que o j-ésimo elemento de $\hat{\mathbf{p}}$ é

[1] Uma matriz N x M é de posto completo quando r = min (M,N)

$$\hat{p}_j = \sum_{i=1}^{M} \frac{\beta_i}{S_i} v_{ij}$$

Análise da Unicidade e Estabilidade do estimador de Mínimos Quadrados usando a SVD como ferramenta:

A análise do estimador de Mínimos Quadrados via decomposição em valores singulares considerando a que a matriz de sensibilidade é de posto completo e a condição que N ≥ M resultou em

$$\overline{\hat{\mathbf{p}}} = \sum_{i=1}^{M} \overline{\mathbf{v}}_i \frac{\beta_i}{S_i}$$

Veja que para a garantia da unicidade da solução estimada $\hat{\mathbf{p}}$ temos que ter, obrigatoriamente, M valores singulares diferentes de zero, caso contrário, teremos infinitas possíveis soluções produzindo o mesmo ajuste (mais adiante provaremos esta informação matematicamente).

Note no entanto que a condição de N ≥ M = r, garante a unicidade da solução $\hat{\mathbf{p}}$ (MQ) porém a simples existência de M valores singulares diferentes de zero não garante a estabilidade da solução. Veja que o j-ésimo elemento de $\hat{\mathbf{p}}$ é

$$\hat{p}_j = \sum_{i=1}^{M} \frac{\beta_i}{S_i} v_{ij}$$

Note que a existência de valores singulares muito próximo a zero causará uma amplificação do ruído dos dados que está presente nos elementos do vetor $\overline{\boldsymbol{\beta}}$ uma vez que este vetor é $\overline{\boldsymbol{\beta}} = \overline{\overline{\mathbf{U}}}^T \overline{\mathbf{y}}^o$ [2]

[2] $\overline{\boldsymbol{\beta}} = \overline{\overline{\mathbf{U}}}^T \overline{\mathbf{y}}^o$ é a projeção dos dados observados nos autovetores da matriz $\overline{\overline{\mathbf{U}}}$

Esta análise da estabilidade de $\hat{\mathbf{p}}$ usando a SVD pode também ser feita via a matriz de covariancia de $\hat{\mathbf{p}}$

$$\text{COV}\left\{\hat{\mathbf{p}}\right\} = \sigma^2 \left(\overline{\overline{\mathbf{A}}}^{\mathrm{T}}\overline{\overline{\mathbf{A}}}\right)^{-1}$$

Sob as mesmas condições N ≥ M = r que estabelecemos para a garantia da unicidade da solução MQ, a decomposição em valores singulares da matriz de sensibilidade é igual a $\overline{\overline{\mathbf{A}}} = \overline{\overline{\mathbf{U}}}_{\mathbf{r}}\,\overline{\overline{\mathbf{S}}}_{\mathbf{r}}\,\overline{\overline{\mathbf{V}}}_{\mathbf{r}}^{\mathbf{T}}$ em que r = M, logo temos que

$$\text{COV}\left\{\hat{\mathbf{p}}\right\} = \sigma^2 \left(\overline{\overline{\mathbf{V}}}_{\mathrm{M}}\,\overline{\overline{\mathbf{S}}}_{\mathrm{M}}\,\overline{\overline{\mathbf{U}}}_{\mathrm{M}}^{\mathrm{T}}\,\overline{\overline{\mathbf{U}}}_{\mathrm{M}}\,\overline{\overline{\mathbf{S}}}_{\mathrm{M}}\,\overline{\overline{\mathbf{V}}}_{\mathrm{M}}^{\mathrm{T}}\right)^{-1}$$

$$\text{COV}\left\{\hat{\mathbf{p}}\right\} = \sigma^2 \left(\overline{\overline{\mathbf{V}}}_{\mathrm{M}}\,\overline{\overline{\mathbf{S}}}_{\mathrm{M}}^{2}\,\overline{\overline{\mathbf{V}}}_{\mathrm{M}}^{\mathrm{T}}\right)^{-1}$$

como r = m temos que $\overline{\overline{\mathbf{V}}}_{M}^{-1} = \overline{\overline{\mathbf{V}}}_{M}^{\mathrm{T}}$. Então a matriz de covariância do estiamdor MQ (N ≥ M = r) é

$$\text{COV}\left\{\hat{\mathbf{p}}\right\} = \sigma^2\,\overline{\overline{\mathbf{V}}}_{\mathrm{M}}\,\overline{\overline{\mathbf{S}}}_{\mathrm{M}}^{-2}\,\overline{\overline{\mathbf{V}}}_{\mathrm{M}}^{\mathrm{T}}$$

Veja que a covariância do kj-ésimo elemento da matriz $\text{cov}\left\{\hat{\mathbf{p}}\right\}$ é

$$\text{cov}\left\{\hat{\mathbf{p}}_{\text{kj}}\right\} = \sigma^2 \sum_{j=1}^{M} \frac{v_{kj}\,v_{kj}}{S_j^{\,2}} \qquad k = 1, \ldots, M$$

Como a variância de $\hat{\mathbf{p}}$ são os elementos da diagonal da matriz de covariância ($\text{cov}\left\{\hat{\mathbf{p}}\right\}$) então temos que a variância do k-ésimo parâmetro é

$$\text{var}\left\{\hat{\mathbf{p}}_{\text{k}}\right\} = \sigma^2 \sum_{j=1}^{M} \frac{v_{kj}^{2}}{S_j^{\,2}}$$

Note que a variância do k-ésimo parâmetro estimado via estimador MQ (N≥M= r) é inversamente proporcional aos valores singulares da matriz de sensibilidade $\overline{\overline{\mathbf{A}}}$. Os elementos da matriz de autovetores (v_{kj}) não exercem influência na variância dos parâmetros pois $\sum_{j=1}^{M} v_{kj}^{2} = 1$. Concluímos, também, através da matriz de covariância de $\hat{\bar{\mathbf{p}}}$ que a presença de valores singulares próximos a zero causam a instabilidade da solução estimada via estimador MQ pois haverá uma amplificação do ruído contido nos dados (variável σ^2 que é a variância do ruído dos dados).

Análise do ajuste dos dados observados produzido pela solução estimada via Mínimos Quadrados

Vamos analisar o ajuste dos dados observados produzido pela solução estimada via estimador MQ. Considerando que os dados observados estão contaminados por ruído aleatório aditivo $\bar{\boldsymbol{\varepsilon}}$ e que em um problema linear a componente determinística (vetor dos dados ajustados ou calculados) é $\bar{\mathbf{y}}^{C} = \bar{\bar{\mathbf{A}}}\,\bar{\mathbf{p}}$. Então os dados geofísicos observados $\bar{\mathbf{y}}^{\mathbf{o}}$ pode ser expresso por

$$\bar{\mathbf{y}}^{\mathbf{o}} = \bar{\mathbf{y}}^{c} + \bar{\boldsymbol{\varepsilon}}$$

$$\bar{\mathbf{y}}^{\mathbf{o}} = \bar{\bar{\mathbf{A}}}\,\bar{\mathbf{p}} + \bar{\boldsymbol{\varepsilon}}$$

Como a estimativa do vetor de parâmetros via estimador MQ é dado por

$$\hat{\bar{\mathbf{p}}} = \left(\bar{\bar{\mathbf{A}}}^{\mathbf{T}}\bar{\bar{\mathbf{A}}}\right)^{\mathbf{-1}}\bar{\bar{\mathbf{A}}}^{\mathbf{T}}\bar{\mathbf{y}}^{\mathbf{o}},$$

então o ajuste o ajuste dos dados observados produzido pela solução estimada via estimador MQ é um AJUSTE dado por

$$\bar{\mathbf{y}}^{\mathbf{o}} = \bar{\bar{\mathbf{A}}}\,\hat{\bar{\mathbf{p}}} + \bar{\boldsymbol{\varepsilon}}$$

$$\bar{\mathbf{y}}^{\mathbf{o}} = \bar{\bar{\mathbf{A}}}\left(\bar{\bar{\mathbf{A}}}^{T}\bar{\bar{\mathbf{A}}}\right)^{-1}\bar{\bar{\mathbf{A}}}^{\mathbf{T}}\bar{\mathbf{y}}^{\mathbf{o}} + \bar{\boldsymbol{\varepsilon}}$$

$$\bar{\boldsymbol{\varepsilon}} = \bar{\mathbf{y}}^{\mathbf{o}} - \bar{\bar{\mathbf{A}}}\left(\bar{\bar{\mathbf{A}}}^{T}\bar{\bar{\mathbf{A}}}\right)^{-1}\bar{\bar{\mathbf{A}}}^{\mathbf{T}}\bar{\mathbf{y}}^{\mathbf{o}}$$

$$\bar{\boldsymbol{\varepsilon}} = \left[\bar{\bar{\mathbf{I}}} - \bar{\bar{\mathbf{A}}} \left(\bar{\bar{\mathbf{A}}}^{T} \bar{\bar{\mathbf{A}}} \right)^{-1} \bar{\bar{\mathbf{A}}}^{\mathbf{T}} \right] \bar{\mathbf{y}}^{\mathbf{o}}$$

Fazendo a decomposição em valores singulares e lembrando que r = M para o estimador MQ temos que o resíduo é dado por

$$\bar{\boldsymbol{\varepsilon}} = \left[\bar{\bar{\mathbf{I}}} - \bar{\bar{\mathbf{U}}}_{\mathrm{M}} \bar{\bar{\mathbf{S}}}_{\mathrm{M}} \bar{\bar{\mathbf{V}}}_{\mathrm{M}}^{\mathrm{T}} \left(\bar{\bar{\mathbf{V}}}_{\mathrm{M}} \bar{\bar{\mathbf{S}}}_{\mathrm{M}} \bar{\bar{\mathbf{U}}}_{\mathrm{M}}^{\mathrm{T}} \bar{\bar{\mathbf{U}}}_{\mathrm{M}} \bar{\bar{\mathbf{S}}}_{\mathrm{M}} \bar{\bar{\mathbf{V}}}_{\mathrm{M}}^{\mathrm{T}} \right)^{-1} \bar{\bar{\mathbf{V}}}_{\mathrm{M}} \bar{\bar{\mathbf{S}}}_{\mathrm{M}} \bar{\bar{\mathbf{U}}}_{\mathrm{M}}^{\mathrm{T}} \right] \bar{\mathbf{y}}^{\mathbf{o}}$$

Como $\bar{\bar{\mathbf{U}}}_{\mathrm{M}}^{\mathrm{T}} \bar{\bar{\mathbf{U}}}_{\mathrm{M}} = \bar{\bar{\mathrm{I}}}_{\mathrm{M}}$ temos que

$$\bar{\boldsymbol{\varepsilon}} = \left[\bar{\bar{\mathbf{I}}} - \bar{\bar{\mathbf{U}}}_{\mathrm{M}} \bar{\bar{\mathbf{S}}}_{\mathrm{M}} \bar{\bar{\mathbf{V}}}_{\mathrm{M}}^{\mathrm{T}} \left(\bar{\bar{\mathbf{V}}}_{\mathrm{M}} \bar{\bar{\mathbf{S}}}_{\mathrm{M}}^{2} \bar{\bar{\mathbf{V}}}_{\mathrm{M}}^{\mathrm{T}} \right)^{-1} \bar{\bar{\mathbf{V}}}_{\mathrm{M}} \bar{\bar{\mathbf{S}}}_{\mathrm{M}} \bar{\bar{\mathbf{U}}}_{\mathrm{M}}^{\mathrm{T}} \right] \bar{\mathbf{y}}^{\mathbf{o}}$$

Como no estimador MQ o posto da matriz $\bar{\bar{\mathbf{A}}}$ é igual ao número de parâmetros (i.e., r = M) temos que $\bar{\bar{\mathbf{V}}}_{\mathrm{M}}^{-1} = \bar{\bar{\mathbf{V}}}_{\mathrm{M}}^{T}$. Usando a propriedade da inversa do produto de várias matrizes é o produto das inversas na ordem inversa [i.e. $\left(\bar{\bar{\mathbf{A}}} \bar{\bar{\mathbf{C}}} \bar{\bar{\mathbf{B}}} \right)^{-1} = \bar{\bar{\mathbf{B}}}^{-1} \bar{\bar{\mathbf{C}}}^{-1} \bar{\bar{\mathbf{A}}}^{-1}$] então o termo a ser invertido na equação acima

$$\begin{aligned} \left(\bar{\bar{\mathbf{V}}}_{\mathrm{M}} \bar{\bar{\mathbf{S}}}_{\mathrm{M}}^{2} \bar{\bar{\mathbf{V}}}_{\mathrm{M}}^{\mathrm{T}} \right)^{-1} &= \left(\bar{\bar{\mathbf{V}}}_{\mathrm{M}}^{\mathrm{T}} \right)^{-1} \bar{\bar{\mathbf{S}}}_{\mathrm{M}}^{-2} \left(\bar{\bar{\mathbf{V}}}_{\mathrm{M}} \right)^{-1} \\ &= \bar{\bar{\mathbf{V}}}_{\mathrm{M}} \bar{\bar{\mathbf{S}}}_{\mathrm{M}}^{-2} \bar{\bar{\mathbf{V}}}_{\mathrm{M}}^{\mathrm{T}} \end{aligned}$$

Logo o vetor de resíduos do estimador MQ pode ser escrito como

$$\bar{\boldsymbol{\varepsilon}} = \left[\bar{\bar{\mathbf{I}}} - \bar{\bar{\mathbf{U}}}_{\mathrm{M}} \bar{\bar{\mathbf{S}}}_{\mathrm{M}} \bar{\bar{\mathbf{V}}}_{\mathrm{M}}^{\mathrm{T}} \bar{\bar{\mathbf{V}}}_{\mathrm{M}} \bar{\bar{\mathbf{S}}}_{\mathrm{M}}^{-2} \bar{\bar{\mathbf{V}}}_{\mathrm{M}}^{T} \bar{\bar{\mathbf{V}}}_{\mathrm{M}} \bar{\bar{\mathbf{S}}}_{\mathrm{M}} \bar{\bar{\mathbf{U}}}_{\mathrm{M}}^{\mathrm{T}} \right] \bar{\mathbf{y}}^{\mathbf{o}}$$

Como $\bar{\bar{\mathbf{V}}}_{\mathrm{M}}^{\mathrm{T}} \bar{\bar{\mathbf{V}}}_{\mathrm{M}} = \bar{\bar{\mathrm{I}}}_{\mathrm{M}}$ temos que o resíduo é dado por

$$\bar{\boldsymbol{\varepsilon}} = \left[\bar{\bar{\mathbf{I}}} - \bar{\bar{\mathbf{U}}}_{\mathrm{M}} \bar{\bar{\mathbf{S}}}_{\mathrm{M}} \bar{\bar{\mathbf{S}}}_{\mathrm{M}}^{-2} \bar{\bar{\mathbf{S}}}_{\mathrm{M}} \bar{\bar{\mathbf{U}}}_{\mathrm{M}}^{\mathrm{T}} \right] \bar{\mathbf{y}}^{\mathbf{o}}$$

$$\bar{\boldsymbol{\varepsilon}} = \left[\bar{\bar{\mathbf{I}}} - \bar{\bar{\mathbf{U}}}_{\mathrm{M}} \bar{\bar{\mathbf{U}}}_{\mathrm{M}}^{\mathrm{T}} \right] \bar{\mathbf{y}}^{\mathbf{o}}$$

Como $\overline{\overline{\mathbf{U}}}_r \, \overline{\overline{\mathbf{U}}}_r^{\mathbf{T}} + \overline{\overline{\mathbf{U}}}_{N-r} \, \overline{\overline{\mathbf{U}}}_{N-r}^{\mathbf{T}} = \overline{\overline{\mathbf{I}}}_{(N \times N)}$ e r = M temos que $\overline{\overline{\mathbf{U}}}_M \, \overline{\overline{\mathbf{U}}}_M^{\mathbf{T}} + \overline{\overline{\mathbf{U}}}_{N-M} \, \overline{\overline{\mathbf{U}}}_{N-M}^{\mathbf{T}} = \overline{\overline{\mathbf{I}}}_{(N \times N)}$ portanto $\overline{\overline{\mathbf{U}}}_{N-M} \, \overline{\overline{\mathbf{U}}}_{N-M}^{\mathbf{T}} = \overline{\overline{\mathbf{I}}}_{(N \times N)} - \overline{\overline{\mathbf{U}}}_M \, \overline{\overline{\mathbf{U}}}_M^{T}$.

Desta forma o vetor de resíduos do estimador MQ pode ser escrito como

$$\overline{\boldsymbol{\varepsilon}} = \overline{\overline{\mathbf{U}}}_{\mathrm{N-M}} \, \overline{\overline{\mathbf{U}}}_{\mathrm{N-M}}^{\mathrm{T}} \, \overline{\mathbf{y}}^{\mathbf{o}}$$

Logo os resíduos do estimador MQ é a projeção dos dados observados no espaço não iluminado das observações.