Medidas de Eficiência dos Estimadores Genéricos

Obtemos as duas equações dos estimadores gerais $\tilde{\bar{\mathbf{p}}}$ que chamaremos de **estimador genérico** são:

$$\tilde{\bar{\mathbf{p}}} = \bar{\mathbf{p}}^{\mathbf{o}} + \left(\bar{\bar{\mathbf{A}}}^{\mathrm{T}} \bar{\bar{\mathbf{W}}}_y \bar{\bar{\mathbf{A}}} + \mu(\delta) \bar{\bar{\mathbf{W}}}_p\right)^{-1} \bar{\bar{\mathbf{A}}}^{\mathrm{T}} \bar{\bar{\mathbf{W}}}_y \left(\bar{\mathbf{y}}^{\mathbf{o}} - \bar{\bar{\mathbf{A}}}\, \bar{\mathbf{p}}^{\mathbf{o}}\right) \tag{1}$$

e

$$\tilde{\bar{\mathbf{p}}} = \bar{\mathbf{p}}^{\mathbf{o}} + \bar{\bar{\mathbf{W}}}_p^{-1} \bar{\bar{\mathbf{A}}}^{\mathrm{T}} \left[\bar{\bar{\mathbf{A}}} \bar{\bar{\mathbf{W}}}_p^{-1} \bar{\bar{\mathbf{A}}}^{\mathrm{T}} + \mu(\delta) \bar{\bar{\mathbf{W}}}_y^{-1}\right]^{-1} \left(\bar{\mathbf{y}}^{\mathbf{o}} - \bar{\bar{\mathbf{A}}}\, \bar{\mathbf{p}}^{\mathbf{o}}\right) \tag{2}$$

Dentre os possíveis medidores de eficiência e desempenho dos estimadores $\tilde{\bar{\mathbf{p}}}$ consideraremos:

1. A esperança de $\tilde{\bar{\mathbf{p}}}$;
2. A matriz de covariância de $\tilde{\bar{\mathbf{p}}}$;
3. A esperança da distância $\left(\tilde{\bar{\mathbf{p}}} - \bar{\mathbf{p}}\right)^{\mathbf{T}} \left(\tilde{\bar{\mathbf{p}}} - \bar{\mathbf{p}}\right)$
4. A matriz de resolução associada a $\tilde{\bar{\mathbf{p}}}$; e
5. A matriz de densidade de informação associada a $\tilde{\bar{\mathbf{p}}}$;

Os estimadores genérico (1) e (2) podem ser escritos na forma geral

$$\boxed{\tilde{\bar{\mathbf{p}}} = \bar{\mathbf{p}}^{\mathbf{o}} + \bar{\bar{\mathbf{H}}}\left(\bar{\mathbf{y}}^{\mathbf{o}} - \bar{\bar{\mathbf{A}}}\, \bar{\mathbf{p}}^{\mathbf{o}}\right)} \tag{3}$$

com

$$\bar{\bar{\mathbf{H}}} = \left(\bar{\bar{\mathbf{A}}}^{\mathrm{T}} \bar{\bar{\mathbf{W}}}_y \bar{\bar{\mathbf{A}}} + \mu(\delta) \bar{\bar{\mathbf{W}}}_p\right)^{-1} \bar{\bar{\mathbf{A}}}^{\mathrm{T}} \bar{\bar{\mathbf{W}}}_y \tag{4}$$

no caso do estimador (1) e com

$$\bar{\bar{\mathbf{H}}} = \bar{\bar{\mathbf{W}}}_p^{-1} \bar{\bar{\mathbf{A}}}^{\mathrm{T}} \left[\bar{\bar{\mathbf{A}}} \bar{\bar{\mathbf{W}}}_p^{-1} \bar{\bar{\mathbf{A}}}^{\mathrm{T}} + \mu(\delta) \bar{\bar{\mathbf{W}}}_y^{-1}\right]^{-1} \tag{5}$$

no caso do estimador (2) .

Análise estatística da solução estimada :

(1) Esperança de $\tilde{\overline{\mathbf{p}}}$:

Considerando as premissas estatísticas estabelecidas na Figura 1 e as propriedades da esperança, temos que a esperança dos estimadores genéricos escritos na forma (3) é

$$E\left\{\tilde{\overline{\mathbf{p}}}\right\} = E\left\{\overline{\mathbf{p}}^{\mathbf{o}} + \overline{\overline{\mathbf{H}}}\left(\overline{\mathbf{y}}^{\mathbf{o}} - \overline{\overline{\mathbf{A}}}\,\overline{\mathbf{p}}^{\mathbf{o}}\right)\right\}$$

$$E\left\{\tilde{\overline{\mathbf{p}}}\right\} = E\left\{\overline{\mathbf{p}}^{\mathbf{o}}\right\} + E\left\{\overline{\overline{\mathbf{H}}}\,\overline{\mathbf{y}}^{\mathbf{o}}\right\} - E\left\{\overline{\overline{\mathbf{H}\mathbf{A}}}\,\overline{\mathbf{p}}^{\mathbf{o}}\right\}$$

Esperança de $\tilde{\overline{\mathbf{p}}}$

$$\tilde{\overline{\mathbf{p}}} = \overline{\mathbf{p}}^{\mathbf{o}} + \overline{\overline{\mathbf{H}}}\left(\overline{\mathbf{y}}^{\mathbf{o}} - \overline{\overline{\mathbf{A}}}\,\overline{\mathbf{p}}^{\mathbf{o}}\right)$$

Esperança :

Premissas estatísticas 1 1 — — — — 1 1

Erro aditivo

Erro com média nula

variáveis independentes sem erros

parâmetros não são aleatórios

Figura 1

Usando a premissa 7 que as variáveis independentes não são v.a. temos

$$E\left\{\tilde{\overline{\mathbf{p}}}\right\} = \overline{\mathbf{p}}^{\mathbf{o}} + \overline{\overline{\mathbf{H}}}\; E\left\{\overline{\mathbf{y}}^{\mathbf{o}}\right\} - \overline{\overline{\mathbf{H}\mathbf{A}}}\,\overline{\mathbf{p}}^{\mathbf{o}}$$

$$E\left\{\tilde{\overline{\mathbf{p}}}\right\} = \overline{\overline{\mathbf{H}}}\; E\left\{\overline{\mathbf{y}}^{\mathbf{o}}\right\} + \left(\overline{\overline{\mathbf{I}}} - \overline{\overline{\mathbf{H}\mathbf{A}}}\right)\overline{\mathbf{p}}^{\mathbf{o}}$$

Usando a premissa 1 que os erros são aditivos ($\overline{\mathbf{y}}^{o} = \overline{\mathbf{y}}^{c} + \overline{\boldsymbol{\varepsilon}}$) e usando a informação que em um problema linear a componente determinística (vetor dos dados ajustados ou calculados) é $\overline{\mathbf{y}}^{C} = \overline{\overline{\mathbf{A}}}\,\overline{\mathbf{p}}$ temos que

$$E\left\{\tilde{\overline{\mathbf{p}}}\right\} = \overline{\overline{\mathbf{H}}}\,E\left\{\overline{\overline{\mathbf{A}}}\,\overline{\mathbf{p}} + \overline{\boldsymbol{\varepsilon}}\right\} + \left(\overline{\overline{\mathbf{I}}} - \overline{\overline{\mathbf{H}\mathbf{A}}}\right)\overline{\mathbf{p}}^{\mathbf{o}}$$

$$E\left\{\tilde{\overline{\mathbf{p}}}\right\}=\overline{\overline{\mathbf{H}}}E\left\{\overline{\overline{\mathbf{A}}}\,\overline{\mathbf{p}}\right\}+\overline{\overline{\mathbf{H}}}E\left\{\overline{\boldsymbol{\varepsilon}}\right\}+\left(\overline{\overline{\mathbf{I}}}-\overline{\overline{\mathbf{H}\,\mathbf{A}}}\right)\overline{\mathbf{p}}^{\mathbf{o}}$$

Pela premissa 2 os erros tem média nula o que implica $E[\boldsymbol{\varepsilon}]=\overline{0}$ então

$$E\left\{\tilde{\overline{\mathbf{p}}}\right\}=\overline{\overline{\mathbf{H}}}E\left\{\overline{\overline{\mathbf{A}}}\,\overline{\mathbf{p}}\right\}+\left(\overline{\overline{\mathbf{I}}}-\overline{\overline{\mathbf{H}\,\mathbf{A}}}\right)\overline{\mathbf{p}}^{\mathbf{o}}$$

Usando a premissa 7 que as variáveis independentes não são v.a. temos

$$E\left\{\tilde{\overline{\mathbf{p}}}\right\}=\overline{\overline{\mathbf{H}\,\mathbf{A}}}E\left\{\overline{\mathbf{p}}\right\}+\left(\overline{\overline{\mathbf{I}}}-\overline{\overline{\mathbf{H}\,\mathbf{A}}}\right)\overline{\mathbf{p}}^{\mathbf{o}}$$

Usando a premissa 8 que os parâmetros não são v.a. temos

$$\boxed{E\left\{\tilde{\overline{\mathbf{p}}}\right\}=\overline{\overline{\mathbf{H}\,\mathbf{A}}}\,\overline{\mathbf{p}}+\left(\overline{\overline{\mathbf{I}}}-\overline{\overline{\mathbf{H}\,\mathbf{A}}}\right)\overline{\mathbf{p}}^{\mathbf{o}}} \qquad (6)$$

(2) Matriz de covariância de $\tilde{\overline{\mathbf{p}}}$:

Considerando as premissas estatísticas estabelecidas na Figura 2 e as propriedades da variância, temos que

$$\text{COV}\left\{\hat{\overline{\mathbf{p}}}\right\}=E\left\{\left[\hat{\overline{\mathbf{p}}}-E(\hat{\overline{\mathbf{p}}})\right]\left[\hat{\overline{\mathbf{p}}}-E(\hat{\overline{\mathbf{p}}})\right]^{T}\right\}$$

Matriz de Covariância $\tilde{\overline{\mathbf{p}}}$

$$\tilde{\overline{\mathbf{p}}}=\overline{\mathbf{p}}^{\mathbf{o}}+\overline{\overline{\mathbf{H}}}\left(\overline{\mathbf{y}}^{\mathbf{o}}-\overline{\overline{\mathbf{A}}}\,\overline{\mathbf{p}}^{\mathbf{o}}\right)$$

Premissas estatísticas 1 1 — — — — 1 1

Erro aditivo
Erro com média nula

variáveis independentes sem erros
parâmetros não são aleatórios

$$\text{COV}\left\{\tilde{\overline{\mathbf{p}}}\right\}=E\left\{\left[\tilde{\overline{\mathbf{p}}}-E\{\tilde{\overline{\mathbf{p}}}\}\right]\left[\tilde{\overline{\mathbf{p}}}-E\{\tilde{\overline{\mathbf{p}}}\}\right]^{T}\right\}$$

Figura 2

como o estimador genérico [equação (3)] é

$$\tilde{\overline{\mathbf{p}}}=\overline{\mathbf{p}}^{\mathbf{o}}+\overline{\overline{\mathbf{H}}}\left(\overline{\mathbf{y}}^{\mathbf{o}}-\overline{\overline{\mathbf{A}}}\,\overline{\mathbf{p}}^{\mathbf{o}}\right)$$

e, pela equação (6), a esperança de $\tilde{\bar{\mathbf{p}}}$

$$E\left\{\tilde{\bar{\mathbf{p}}}\right\} = \overline{\overline{\mathbf{H}\,\mathbf{A}}}\,\bar{\mathbf{p}} + \left(\overline{\overline{\mathbf{I}}} - \overline{\overline{\mathbf{H}\,\mathbf{A}}}\right)\bar{\mathbf{p}}^{\mathbf{o}}$$

então temos que

$$\text{COV}\left\{\tilde{\bar{\mathbf{p}}}\right\} = E\left\{\left[\tilde{\bar{\mathbf{p}}} - E\left(\tilde{\bar{\mathbf{p}}}\right)\right]\left[\tilde{\bar{\mathbf{p}}} - E\left(\tilde{\bar{\mathbf{p}}}\right)\right]^{\mathbf{T}}\right\}$$

$$\text{COV}\left\{\tilde{\bar{\mathbf{p}}}\right\} = E\left\{\left[\bar{\mathbf{p}}^{\mathbf{o}} + \overline{\overline{\mathbf{H}}}\left(\bar{\mathbf{y}}^{\mathbf{o}} - \overline{\overline{\mathbf{A}}}\,\bar{\mathbf{p}}^{\mathbf{o}}\right) - \overline{\overline{\mathbf{H}\,\mathbf{A}}}\,\bar{\mathbf{p}} - \left(\overline{\overline{\mathbf{I}}} - \overline{\overline{\mathbf{H}\,\mathbf{A}}}\right)\bar{\mathbf{p}}^{\mathbf{o}}\right]\right.$$
$$\left.\left[\bar{\mathbf{p}}^{\mathbf{o}} + \overline{\overline{\mathbf{H}}}\left(\bar{\mathbf{y}}^{\mathbf{o}} - \overline{\overline{\mathbf{A}}}\,\bar{\mathbf{p}}^{\mathbf{o}}\right) - \overline{\overline{\mathbf{H}\,\mathbf{A}}}\,\bar{\mathbf{p}} - \left(\overline{\overline{\mathbf{I}}} - \overline{\overline{\mathbf{H}\,\mathbf{A}}}\right)\bar{\mathbf{p}}^{\mathbf{o}}\right]^{\mathbf{T}}\right\}$$

$$\text{COV}\left\{\tilde{\bar{\mathbf{p}}}\right\} = E\left\{\left[\cancel{\bar{\mathbf{p}}^{\mathbf{o}}} + \overline{\overline{\mathbf{H}}}\,\bar{\mathbf{y}}^{\mathbf{o}} - \cancel{\overline{\overline{\mathbf{H}\,\mathbf{A}}}\,\bar{\mathbf{p}}^{\mathbf{o}}} - \overline{\overline{\mathbf{H}\,\mathbf{A}}}\,\bar{\mathbf{p}} - \cancel{\bar{\mathbf{p}}^{\mathbf{o}}} + \cancel{\overline{\overline{\mathbf{H}\,\mathbf{A}}}\,\bar{\mathbf{p}}^{\mathbf{o}}}\right]\right.$$
$$\left.\left[\cancel{\bar{\mathbf{p}}^{\mathbf{o}}} + \overline{\overline{\mathbf{H}}}\,\bar{\mathbf{y}}^{\mathbf{o}} - \cancel{\overline{\overline{\mathbf{H}\,\mathbf{A}}}\,\bar{\mathbf{p}}^{\mathbf{o}}} - \overline{\overline{\mathbf{H}\,\mathbf{A}}}\,\bar{\mathbf{p}} - \cancel{\bar{\mathbf{p}}^{\mathbf{o}}} + \cancel{\overline{\overline{\mathbf{H}\,\mathbf{A}}}\,\bar{\mathbf{p}}^{\mathbf{o}}}\right]^{\mathbf{T}}\right\}$$

$$\text{COV}\left\{\tilde{\bar{\mathbf{p}}}\right\} = E\left\{\left[\overline{\overline{\mathbf{H}}}\,\bar{\mathbf{y}}^{\mathbf{o}} - \overline{\overline{\mathbf{H}\,\mathbf{A}}}\,\bar{\mathbf{p}}\right]\left[\overline{\overline{\mathbf{H}}}\,\bar{\mathbf{y}}^{\mathbf{o}} - \overline{\overline{\mathbf{H}\,\mathbf{A}}}\,\bar{\mathbf{p}}\right]^{\mathbf{T}}\right\}$$

$$\text{COV}\left\{\tilde{\bar{\mathbf{p}}}\right\} = E\left\{\left[\overline{\overline{\mathbf{H}}}\left(\bar{\mathbf{y}}^{\mathbf{o}} - \overline{\overline{\mathbf{A}}}\,\bar{\mathbf{p}}\right)\right]\left[\overline{\overline{\mathbf{H}}}\left(\bar{\mathbf{y}}^{\mathbf{o}} - \overline{\overline{\mathbf{A}}}\,\bar{\mathbf{p}}\right)\right]^{\mathbf{T}}\right\}$$

$$\text{COV}\left\{\tilde{\bar{\mathbf{p}}}\right\} = E\left\{\left[\overline{\overline{\mathbf{H}}}\,\bar{\boldsymbol{\varepsilon}}\right]\left[\overline{\overline{\mathbf{H}}}\,\bar{\boldsymbol{\varepsilon}}\right]^{\mathbf{T}}\right\}$$

$$\text{COV}\left\{\tilde{\bar{\mathbf{p}}}\right\} = E\left\{\overline{\overline{\mathbf{H}}}\,\bar{\boldsymbol{\varepsilon}}\,\bar{\boldsymbol{\varepsilon}}^{\mathbf{T}}\,\overline{\overline{\mathbf{H}}}^{\mathbf{T}}\right\}$$

Usando a premissa 7 que as variáveis independentes não são v.a. temos

$$\text{COV}\left\{\tilde{\bar{\mathbf{p}}}\right\} = \overline{\overline{\mathbf{H}}}\,E\left\{\bar{\boldsymbol{\varepsilon}}\,\bar{\boldsymbol{\varepsilon}}^{\mathbf{T}}\right\}\,\overline{\overline{\mathbf{H}}}^{\mathbf{T}}$$

A esperança de $E\left\{\bar{\boldsymbol{\varepsilon}}\bar{\boldsymbol{\varepsilon}}^{T}\right\}$ que nada mais é que a esperança de uma forma quadrática

(jávimos no Apêndice Matriz de Covariância dos Erros) que

$$E\left\{\bar{\boldsymbol{\varepsilon}}\,\bar{\boldsymbol{\varepsilon}}^T\right\} = \text{cov}(\bar{\boldsymbol{\varepsilon}}) + E(\bar{\boldsymbol{\varepsilon}})\,E(\bar{\boldsymbol{\varepsilon}}^T)$$

Pela premissa 2 os erros tem média nula o que implica $E[\bar{\boldsymbol{\varepsilon}}] = \bar{0}$ temos que

$$E\left\{\bar{\boldsymbol{\varepsilon}}\,\bar{\boldsymbol{\varepsilon}}^T\right\} = \text{cov}(\bar{\boldsymbol{\varepsilon}})$$

Logo a covariância do estimador genérico é

$$\boxed{\text{COV}\left\{\tilde{\bar{\mathbf{p}}}\right\} = \bar{\bar{\mathbf{H}}}\ \text{cov}\left\{\bar{\boldsymbol{\varepsilon}}\right\}\ \bar{\bar{\mathbf{H}}}^{\mathrm{T}}} \qquad (7)$$

(3) Esperança da distância entre $\tilde{\bar{\mathbf{p}}}$ e $\bar{\mathbf{p}}$:

A distância entre $\tilde{\bar{\mathbf{p}}}$ e $\bar{\mathbf{p}}$ pode ser expressa por

$$\|\tilde{\bar{\mathbf{p}}} - \bar{\mathbf{p}}\|_2^2 = (\tilde{\bar{\mathbf{p}}} - \bar{\mathbf{p}})^{\mathrm{T}}(\tilde{\bar{\mathbf{p}}} - \bar{\mathbf{p}})$$

Vamos então calcular a esperança da distância entre $\tilde{\bar{\mathbf{p}}}$ e $\bar{\mathbf{p}}$ i.e.,

$$E\{[\tilde{\bar{\mathbf{p}}} - \bar{\mathbf{p}}]^{\mathrm{T}}[\tilde{\bar{\mathbf{p}}} - \bar{\mathbf{p}}]\}$$

Veja que temos a esperança de uma forma quadrática $E\{\bar{\mathbf{x}}^{\mathrm{T}}\bar{\mathbf{x}}\}$ é

$$E\{\bar{\mathbf{x}}^{\mathrm{T}}\bar{\mathbf{x}}\} = Tr\{\text{cov}(\bar{\mathbf{x}})\} + E\{\bar{\mathbf{x}}^T\}\,E\{\bar{\mathbf{x}}^T\}$$

isto porque

$$\text{cov}\{\bar{\mathbf{x}}\} = E\{[\bar{\mathbf{x}} - E\{\bar{\mathbf{x}}\}][\bar{\mathbf{x}} - E\{\bar{\mathbf{x}}\}]^T\}$$

$$\text{cov}\{\bar{\mathbf{x}}\} = E\{\bar{\mathbf{x}}\,\bar{\mathbf{x}}^T\} - 2E\{\bar{\mathbf{x}}\}E\{\bar{\mathbf{x}}\}^T + E\{\bar{\mathbf{x}}\}E\{\bar{\mathbf{x}}\}^T$$
$$\text{cov}\{\bar{\mathbf{x}}\} = E\{\bar{\mathbf{x}}\,\bar{\mathbf{x}}^T\} - E\{\bar{\mathbf{x}}\}E\{\bar{\mathbf{x}}\}^T$$

$$E\{\bar{\mathbf{x}}\,\bar{\mathbf{x}}^T\} = \text{cov}\{\bar{\mathbf{x}}\} + E\{\bar{\mathbf{x}}\}E\{\bar{\mathbf{x}}\}^T$$

Usando o operador traço temos:

$$tr\left\{E\{\overline{\mathbf{x}}\,\overline{\mathbf{x}}^T\}\right\} = tr\left\{\mathrm{cov}\{\overline{\mathbf{x}}\}\right\} + tr\left\{E\{\overline{\mathbf{x}}\}E\{\overline{\mathbf{x}}\}^T\right\}$$

$$E\{\overline{\mathbf{x}}^T\overline{\mathbf{x}}\} = tr\left\{\mathrm{cov}\{\overline{\mathbf{x}}\}\right\} + E\{\overline{\mathbf{x}}^T\}E\{\overline{\mathbf{x}}\}$$

Então a

$$E\{[\tilde{\overline{\mathbf{p}}} - \overline{\mathbf{p}}]^{\mathrm{T}}[\tilde{\overline{\mathbf{p}}} - \overline{\mathbf{p}}]\} = tr\{\mathrm{cov}(\tilde{\overline{\mathbf{p}}} - \overline{\mathbf{p}})\} + E\{(\tilde{\overline{\mathbf{p}}} - \overline{\mathbf{p}})^T\}E\{(\tilde{\overline{\mathbf{p}}} - \overline{\mathbf{p}})\}$$

Vamos calcular o termo

$$E\{(\tilde{\overline{\mathbf{p}}} - \overline{\mathbf{p}})\} = E\{\tilde{\overline{\mathbf{p}}}\} - E\{\overline{\mathbf{p}}\}$$

$$E\{(\tilde{\overline{\mathbf{p}}} - \overline{\mathbf{p}})\} = E\{\tilde{\overline{\mathbf{p}}}\} - \overline{\mathbf{p}}$$

Substituindo a equação (6) temos

$$E\{(\tilde{\overline{\mathbf{p}}} - \overline{\mathbf{p}})\} = \overline{\overline{\mathbf{HA}}}\,\overline{\mathbf{p}} + \left(\overline{\overline{\mathbf{I}}} - \overline{\overline{\mathbf{HA}}}\right)\overline{\mathbf{p}}^{\mathbf{o}} - \overline{\mathbf{p}}$$

$$E\{(\tilde{\overline{\mathbf{p}}} - \overline{\mathbf{p}})\} = \left(\overline{\overline{\mathbf{HA}}} - \overline{\overline{\mathbf{I}}}\right)\overline{\mathbf{p}} + \left(\overline{\overline{\mathbf{I}}} - \overline{\overline{\mathbf{HA}}}\right)\overline{\mathbf{p}}^{\mathbf{o}}$$

$$E\{(\tilde{\overline{\mathbf{p}}} - \overline{\mathbf{p}})\} = \left(\overline{\overline{\mathbf{HA}}} - \overline{\overline{\mathbf{I}}}\right)(\overline{\mathbf{p}} - \overline{\mathbf{p}}^{\mathbf{o}})$$

Substituindo a equação acima na expressão da Esperança da distância temos

$$E\{[\tilde{\overline{\mathbf{p}}} - \overline{\mathbf{p}}]^{\mathrm{T}}[\tilde{\overline{\mathbf{p}}} - \overline{\mathbf{p}}]\} = tr\{\mathrm{cov}(\tilde{\overline{\mathbf{p}}} - \overline{\mathbf{p}})\} + (\overline{\mathbf{p}} - \overline{\mathbf{p}}^{\mathbf{o}})^{\mathrm{T}}\left(\overline{\overline{\mathbf{HA}}} - \overline{\overline{\mathbf{I}}}\right)^T\left(\overline{\overline{\mathbf{HA}}} - \overline{\overline{\mathbf{I}}}\right)(\overline{\mathbf{p}} - \overline{\mathbf{p}}^{\mathbf{o}})$$

Como a covariância de uma constante é zero então a $\mathrm{cov}(\tilde{\overline{\mathbf{p}}} - \overline{\mathbf{p}}) = \mathrm{cov}(\tilde{\overline{\mathbf{p}}})$.

Substituindo a equação (7) na equação acima temos

$$E\{[\tilde{\bar{\mathbf{p}}} - \bar{\mathbf{p}}]^{\mathrm{T}}[\tilde{\bar{\mathbf{p}}} - \bar{\mathbf{p}}]\} = tr\{\mathrm{cov}(\tilde{\bar{\mathbf{p}}})\} + (\bar{\mathbf{p}} - \bar{\mathbf{p}}^{\mathbf{o}})^{\mathrm{T}}\left(\overline{\overline{\mathbf{H}\,\mathbf{A}}} - \bar{\bar{\mathbf{I}}}\right)^{T}\left(\overline{\overline{\mathbf{H}\,\mathbf{A}}} - \bar{\bar{\mathbf{I}}}\right)(\bar{\mathbf{p}} - \bar{\mathbf{p}}^{\mathbf{o}}) \quad (8)$$

Da equação (6) temos que

$$\mathrm{E}\left\{\tilde{\bar{\mathbf{p}}}\right\} = \overline{\overline{\mathbf{H}\,\mathbf{A}}}\,\bar{\mathbf{p}} + \left(\bar{\bar{\mathbf{I}}} - \overline{\overline{\mathbf{H}\,\mathbf{A}}}\right)\bar{\mathbf{p}}^{\mathbf{o}}$$

reescrevendo esta equação temos que

$$\mathrm{E}\left\{\tilde{\bar{\mathbf{p}}}\right\} = \overline{\overline{\mathbf{H}\,\mathbf{A}}}(\bar{\mathbf{p}} - \bar{\mathbf{p}}^{\mathbf{o}}) + \bar{\mathbf{p}}^{\mathbf{o}}$$

subtraindo dos dois lados da equação acima $\bar{\mathbf{p}}$ temos

$$\mathrm{E}\left\{\tilde{\bar{\mathbf{p}}}\right\} - \bar{\mathbf{p}} = \overline{\overline{\mathbf{H}\,\mathbf{A}}}(\bar{\mathbf{p}} - \bar{\mathbf{p}}^{\mathbf{o}}) + \bar{\mathbf{p}}^{\mathbf{o}} - \bar{\mathbf{p}}$$

$$\mathrm{E}\left\{\tilde{\bar{\mathbf{p}}}\right\} - \bar{\mathbf{p}} = \left(\overline{\overline{\mathbf{H}\,\mathbf{A}}} - \bar{\bar{\mathbf{I}}}\right)(\bar{\mathbf{p}} - \bar{\mathbf{p}}^{\mathbf{o}})$$

$$\mathrm{E}\left\{\tilde{\bar{\mathbf{p}}}\right\} - \bar{\mathbf{p}} = \left(\overline{\overline{\mathbf{H}\,\mathbf{A}}} - \bar{\bar{\mathbf{I}}}\right)(\bar{\mathbf{p}} - \bar{\mathbf{p}}^{\mathbf{o}}) \quad (9)$$

Comparando o segundo termo do somatório da equação (8) com a expressão (9), podemos então reescrever a equação (8) como

$$E\{[\tilde{\bar{\mathbf{p}}} - \bar{\mathbf{p}}]^{\mathrm{T}}[\tilde{\bar{\mathbf{p}}} - \bar{\mathbf{p}}]\} = tr\{\mathrm{cov}(\tilde{\bar{\mathbf{p}}})\} + \left[\mathrm{E}\left\{\tilde{\bar{\mathbf{p}}}\right\} - \bar{\mathbf{p}}\right]^{T}\left[\mathrm{E}\left\{\tilde{\bar{\mathbf{p}}}\right\} - \bar{\mathbf{p}}\right]$$

Como $\bar{\mathbf{u}}^{T}\bar{\mathbf{u}} = \|\bar{\mathbf{u}}\|_2^2$ então a equação acima da expressão da Esperança da distância pode ser reescrita como

$$\boxed{E\left\{\left\|\tilde{\bar{\mathbf{p}}} - \bar{\mathbf{p}}\right\|_2^2\right\} = tr\{\mathrm{cov}(\tilde{\bar{\mathbf{p}}})\} + \left\|\mathrm{E}\left\{\tilde{\bar{\mathbf{p}}}\right\} - \bar{\mathbf{p}}\right\|_2^2} \quad (10)$$

Veja que a Esperança da distância $\left\|\tilde{\bar{\mathbf{p}}} - \bar{\mathbf{p}}\right\|_2^2$ equação (10) é formada por duas parcelas:

1) $tr\{\mathrm{cov}(\tilde{\bar{\mathbf{p}}})\}$ é a primeira parcela e representa uma medida de covariância do estimador , ou seja, uma medida da estabilidade da solução estimada

2) $\left\| E\left\{ \tilde{\bar{\mathbf{p}}} \right\} - \bar{\mathbf{p}} \right\|_2^2$ é a segunda parcela representa uma medida da distância da esperança do estimador ($\tilde{\bar{\mathbf{p}}}$) e o vetor de parâmetros ($\bar{\mathbf{p}}$). Veja portanto que esta segunda parcela é uma medida da tendensiosidade do estimador $\tilde{\bar{\mathbf{p}}}$

(4) A matriz de resolução associada a $\tilde{\bar{\mathbf{p}}} - \bar{\mathbf{p}}^{\mathbf{o}}$;

Vamos considerar que os nossos dados geofísicos estão livres de ruído. Imagine que exista um conjunto de parâmetros desconhecido que explique os dados geofísicos sem ruído, ou seja, $\bar{\mathbf{y}}^{\mathbf{o}} = \bar{\bar{\mathbf{A}}}\,\bar{\mathbf{p}}$. Sob estas premissas o estimador genérico (3),

$$\tilde{\bar{\mathbf{p}}} = \bar{\mathbf{p}}^{\mathbf{o}} + \bar{\bar{\mathbf{H}}}\left(\bar{\mathbf{y}}^{\mathbf{o}} - \bar{\bar{\mathbf{A}}}\,\bar{\mathbf{p}}^{\mathbf{o}} \right),$$

pode ser reescrito como

$$\tilde{\bar{\mathbf{p}}} = \bar{\mathbf{p}}^{\mathbf{o}} + \bar{\bar{\mathbf{H}}}\left(\bar{\bar{\mathbf{A}}}\,\bar{\mathbf{p}} - \bar{\bar{\mathbf{A}}}\,\bar{\mathbf{p}}^{\mathbf{o}} \right)$$

$$\tilde{\bar{\mathbf{p}}} = \bar{\mathbf{p}}^{\mathbf{o}} + \overline{\overline{\mathbf{H}\,\mathbf{A}}}\left(\bar{\mathbf{p}} - \bar{\mathbf{p}}^{\mathbf{o}} \right)$$

$$\tilde{\bar{\mathbf{p}}} - \bar{\mathbf{p}}^{\mathbf{o}} = \overline{\overline{\mathbf{H}\,\mathbf{A}}}\left(\bar{\mathbf{p}} - \bar{\mathbf{p}}^{\mathbf{o}} \right) \qquad (11)$$

O produto $\bar{\bar{\mathbf{H}}}\;\bar{\bar{\mathbf{A}}}$ é chamada de matriz resolução de $\tilde{\bar{\mathbf{p}}} - \bar{\mathbf{p}}^{\mathbf{o}}$

$$\boxed{\bar{\bar{\mathbf{R}}} = \bar{\bar{\mathbf{H}}}\;\bar{\bar{\mathbf{A}}}} \qquad (12)$$

A matriz de resolução, $\bar{\bar{\mathbf{R}}}$, é uma matriz de dimensão (M x M).

Note que a matriz de resolução é função apenas da matriz de sensibilidade ($\overline{\overline{\mathbf{A}}}$), do kernel do nosso problema (matriz $\overline{\overline{\mathbf{H}}}$) e das informações a priori adicionadas ao nosso problema (matrizes de peso).

Para entendermos o significado da matriz de resolução vamos considerar que $\overline{\mathbf{p}}^{\mathbf{o}} = \overline{\mathbf{0}}$. Então temos que

$$\tilde{\overline{\mathbf{p}}} = \overline{\overline{\mathbf{H}}}\ \overline{\overline{\mathbf{A}}}\ \overline{\mathbf{p}}$$

$$\tilde{\overline{\mathbf{p}}} = \overline{\overline{\mathbf{R}}}\ \overline{\mathbf{p}} \qquad (13)$$

Veja que a matriz de resolução mede a proximidade entre uma solução particular estimada ($\tilde{\overline{\mathbf{p}}}$) e a solução de parâmetros considerando os dados exatos, $\overline{\mathbf{p}}$, i.e., considerando os dados sem ruído.

Vamos fazer agora um rearranjo da equação (9)

$$\tilde{\overline{\mathbf{p}}} - \overline{\mathbf{p}}^{\mathbf{o}} = \overline{\overline{\mathbf{H}\,\mathbf{A}}}\left(\overline{\mathbf{p}} - \overline{\mathbf{p}}^{\mathbf{o}}\right)$$

$$\tilde{\overline{\mathbf{p}}} = \overline{\mathbf{p}}^{\mathbf{o}} + \overline{\overline{\mathbf{H}\,\mathbf{A}}}\left(\overline{\mathbf{p}} - \overline{\mathbf{p}}^{\mathbf{o}}\right)$$

$$\tilde{\overline{\mathbf{p}}} = \overline{\overline{\mathbf{H}\,\mathbf{A}}}\,\overline{\mathbf{p}} + \left(\overline{\overline{\mathbf{I}}} - \overline{\overline{\mathbf{H}\,\mathbf{A}}}\right)\overline{\mathbf{p}}^{\mathbf{o}} \qquad (14)$$

Veja que se a matriz de resolução $\overline{\overline{\mathbf{R}}} = \overline{\overline{\mathbf{H}}}\ \overline{\overline{\mathbf{A}}}$ for igual a matriz identidade, i.e., $\overline{\overline{\mathbf{R}}} = \overline{\overline{\mathbf{I}}}$ temos que

$$\tilde{\overline{\mathbf{p}}} = \overline{\mathbf{p}}$$

Neste caso em que $\overline{\overline{\mathbf{R}}} = \overline{\overline{\mathbf{I}}}$, então cada elemento do vetor de parâmetros é unicamente determinado, ou seja, os parâmetros podem ser estimados independentemente. Por outro lado, se a matriz de resolução for diferente da matriz identidade, $\overline{\overline{\mathbf{R}}} \neq \overline{\overline{\mathbf{I}}}$ então cada elemento do vetor de parâmetros estimados, $\tilde{\overline{\mathbf{p}}}$, será uma soma ponderada de todos os elementos do vetor de parâmetros $\overline{\mathbf{p}}$. Veja que se considerarmos $\overline{\mathbf{p}}^{\mathbf{o}} = \overline{\mathbf{0}}$ temos que

$$\tilde{\overline{\mathbf{p}}} = \overline{\overline{\mathbf{R}}}\,\overline{\mathbf{p}}$$

$$\begin{pmatrix} \tilde{p}_1 \\ \tilde{p}_2 \\ \vdots \\ \tilde{p}_M \end{pmatrix} = \begin{bmatrix} r_{11} & r_{12} & \cdots & r_{1M} \\ r_{21} & r_{22} & \cdots & r_{2M} \\ \vdots & \vdots & \vdots & \vdots \\ r_{M1} & r_{M2} & \cdots & r_{MM} \end{bmatrix} \begin{pmatrix} p_1 \\ p_2 \\ \vdots \\ p_M \end{pmatrix}$$

$$\tilde{p}_1 = r_{11}\, p_1 + r_{12}\, p_2 + \cdots + r_{1M}\, p_M \;,$$

$$\tilde{p}_2 = r_{21}\, p_1 + r_{22}\, p_2 + \cdots + r_{2M}\, p_M$$

$$\vdots$$

Veja então que se a matriz de resolução for diferente da matriz identidade, $\overline{\overline{\mathbf{R}}} \neq \overline{\overline{\mathbf{I}}}$ o j-ésimo elemento do vetor de parâmetros estimados, $\tilde{\overline{\mathbf{p}}}$ é

$$\begin{pmatrix} \tilde{p}_1 \\ \vdots \\ \tilde{p}_j \\ \vdots \end{pmatrix} = \begin{bmatrix} r_{11} & r_{12} & \cdots & r_{1M} \\ \vdots & \vdots & \cdots & \vdots \\ r_{j1} & r_{j2} & \vdots & r_{jM} \\ \vdots & \vdots & \cdots & \vdots \end{bmatrix} \begin{pmatrix} p_1 \\ \vdots \\ p_j \\ \vdots \end{pmatrix}$$

$$\tilde{p}_j = r_{j1} p_1 + r_{j2} p_2 + \cdots + r_{jM} p_M$$

$$\tilde{p}_j = \sum_{i=1}^{M} r_{ji} \, p_i$$

(5) A matriz densidade de informação (Data resolution matrix) associada a $\tilde{\bar{\mathbf{p}}}$;

Estimamos um vetor de parâmetros $\tilde{\bar{\mathbf{p}}}$

$$\tilde{\bar{\mathbf{p}}} = \bar{\mathbf{p}}^{\mathbf{o}} + \bar{\bar{\mathbf{H}}} \left(\bar{\mathbf{y}}^{\mathbf{o}} - \bar{\bar{\mathbf{A}}} \, \bar{\mathbf{p}}^{\mathbf{o}} \right)$$

Este vetor de parâmetros estimado produz um ajuste (dados geofísicos estimados ou calculados pelo modelo $\tilde{\bar{\mathbf{p}}}$) que é dado por

$$\tilde{\bar{\mathbf{y}}} = \bar{\bar{\mathbf{A}}} \, \tilde{\bar{\mathbf{p}}}$$

$$\tilde{\bar{\mathbf{y}}} = \bar{\bar{\mathbf{A}}} \left[\bar{\mathbf{p}}^{\mathbf{o}} + \bar{\bar{\mathbf{H}}} \left(\bar{\mathbf{y}}^{\mathbf{o}} - \bar{\bar{\mathbf{A}}} \, \bar{\mathbf{p}}^{\mathbf{o}} \right) \right]$$

$$\tilde{\bar{\mathbf{y}}} = \bar{\bar{\mathbf{A}}} \, \bar{\mathbf{p}}^{\mathbf{o}} + \overline{\overline{\mathbf{A} \, \mathbf{H}}} \, \bar{\mathbf{y}}^{\mathbf{o}} - \overline{\overline{\mathbf{A} \, \mathbf{H} \, \mathbf{A}}} \, \bar{\mathbf{p}}^{\mathbf{o}} \qquad (15)$$

Vamos considerar que $\bar{\mathbf{y}}^{a} = \bar{\bar{\mathbf{A}}} \bar{\mathbf{p}}^{\mathbf{o}}$, ou seja, o vetor de parâmetros a priori, $\bar{\mathbf{p}}^{\mathbf{o}}$, produz um ajuste $\bar{\mathbf{y}}^{a}$. Então $\bar{\mathbf{y}}^{a}$ é a resposta geofísica calculada pelo modelo de parâmetros $\bar{\mathbf{p}}^{\mathbf{o}}$. Então a equação (15) pode ser reescrita como

$$\tilde{\bar{\mathbf{y}}} = \bar{\mathbf{y}}^{a} + \overline{\overline{\mathbf{A}}\,\overline{\overline{\mathbf{H}}}}\,\bar{\mathbf{y}}^{\mathbf{o}} - \overline{\overline{\mathbf{A}}\,\overline{\overline{\mathbf{H}}}}\,\bar{\mathbf{y}}^{a}$$

$$\tilde{\bar{\mathbf{y}}} - \bar{\mathbf{y}}^{a} = \overline{\overline{\mathbf{A}}\,\overline{\overline{\mathbf{H}}}}\,(\bar{\mathbf{y}}^{\mathbf{o}} - \bar{\mathbf{y}}^{a}) \qquad (16)$$

O produto $\overline{\overline{\mathbf{A}}}\;\overline{\overline{\mathbf{H}}}$ é chamado de matriz densidade de informação de $\tilde{\bar{\mathbf{y}}} - \bar{\mathbf{y}}^{a}$.

$$\boxed{\underset{(N \times N)}{\overline{\overline{\mathbf{F}}}} = \overline{\overline{\mathbf{A}}}\;\overline{\overline{\mathbf{H}}}} \qquad (17)$$

A matriz densidade de informação, $\overline{\overline{\mathbf{F}}}$, é uma matriz de dimensão (N x N). Note que a matriz densidade de informação é função apenas da matriz de sensibilidade ($\overline{\overline{\mathbf{A}}}$), do kernel do nosso problema (matriz $\overline{\overline{\mathbf{H}}}$) e das informações a priori adicionadas ao nosso problema (matrizes de peso).

Para entendermos o significado da matriz densidade de informação, $\overline{\overline{\mathbf{F}}}$, vamos considerar que $\bar{\mathbf{p}}^{\mathbf{o}} = \bar{\mathbf{0}}$. Então temos que a equação (15),

$$\tilde{\bar{\mathbf{y}}} = \overline{\overline{\mathbf{A}}}\,\bar{\mathbf{p}}^{\mathbf{o}} + \overline{\overline{\mathbf{A}}\,\overline{\overline{\mathbf{H}}}}\,\bar{\mathbf{y}}^{\mathbf{o}} - \overline{\overline{\mathbf{A}}\,\overline{\overline{\mathbf{H}}}\,\overline{\overline{\mathbf{A}}}}\;\bar{\mathbf{p}}^{\mathbf{o}} \;,$$

pode ser reescrita como

$$\tilde{\bar{\mathbf{y}}} = \overline{\overline{\mathbf{A}}\,\overline{\overline{\mathbf{H}}}}\,\bar{\mathbf{y}}^{\mathbf{o}} \qquad (18)$$

Veja que se a matriz densidade de informação, $\overline{\overline{\mathbf{F}}} = \overline{\overline{\mathbf{A}}\,\overline{\overline{\mathbf{H}}}}$, or igual a matriz identidade, i.e., $\overline{\overline{\mathbf{F}}} = \overline{\overline{\mathbf{I}}}$ temos que

$$\tilde{\bar{\mathbf{y}}} = \bar{\mathbf{y}}^{\mathbf{o}}$$

Neste caso em que $\overline{\overline{\mathbf{F}}} = \overline{\overline{\mathbf{I}}}$, então o dado geofísico observado, $\bar{\mathbf{y}}^{\mathbf{o}}$, é exatamente igual ao dado geofísico estimado , $\tilde{\bar{\mathbf{y}}}$. Isto significa dizer que o erro (ruído) dos dados é zero, $\bar{\boldsymbol{\varepsilon}} = \bar{0}$. Veja que a matriz densidade de informação, $\bar{\bar{\mathbf{F}}}$, mede a proximidade entre os dados observados e os dados ajustados. Em outras palavras a matriz densidade de informação, $\bar{\bar{\mathbf{F}}}$, diz o quão próximo os dados estimados $\tilde{\bar{\mathbf{y}}}$ estão dos dados geofísicos observados, $\bar{\mathbf{y}}^{\mathbf{o}}$. Vamos calcular o resíduo $\bar{\boldsymbol{\varepsilon}}$

$$\bar{\boldsymbol{\varepsilon}} = \bar{\mathbf{y}}^{\mathbf{o}} - \tilde{\bar{\mathbf{y}}}$$

Usando a equação que $\tilde{\bar{\mathbf{y}}} = \bar{\mathbf{y}}^{a} + \overline{\overline{\mathbf{A}\,\mathbf{H}}}\,\bar{\mathbf{y}}^{\mathbf{o}} - \overline{\overline{\mathbf{A}\,\mathbf{H}}}\,\bar{\mathbf{y}}^{a}$ temos

$$\bar{\boldsymbol{\varepsilon}} = \bar{\mathbf{y}}^{\mathbf{o}} - \bar{\mathbf{y}}^{a} - \overline{\overline{\mathbf{A}\,\mathbf{H}}}\,\bar{\mathbf{y}}^{\mathbf{o}} + \overline{\overline{\mathbf{A}\,\mathbf{H}}}\,\bar{\mathbf{y}}^{a}$$

$$\bar{\boldsymbol{\varepsilon}} = \left(\bar{\bar{\mathbf{I}}} - \overline{\overline{\mathbf{A}\,\mathbf{H}}}\right)(\bar{\mathbf{y}}^{\mathbf{o}} - \bar{\mathbf{y}}^{a}) \tag{19}$$

A equação (18) mostra explicitamente que se a matriz densidade de informação, $\bar{\bar{\mathbf{F}}} = \overline{\overline{\mathbf{A}\,\mathbf{H}}}$, for igual a matriz identidade, i.e., $\bar{\bar{\mathbf{F}}} = \bar{\bar{\mathbf{I}}}$ temos que o erro (ruído) será zero $\bar{\boldsymbol{\varepsilon}} = \bar{0}$. Isto não é desejável porque os dados contém ruído.

Por outro lado, se a matriz densidade de informação for diferente da matriz identidade, $\bar{\bar{\mathbf{F}}} \neq \bar{\bar{\mathbf{I}}}$ então cada elemento do vetor de dados estimados, $\tilde{\bar{\mathbf{y}}}$, será uma soma ponderada de todos os elementos que compõem o vetor de dados observados $\bar{\mathbf{y}}^{\mathbf{o}}$.

Veja que se considerarmos $\bar{\mathbf{p}}^{\mathbf{o}} = \bar{\mathbf{0}}$ temos que

$$\begin{pmatrix} \tilde{y}_1 \\ \vdots \\ \tilde{y}_i \\ \vdots \\ \tilde{y}_N \end{pmatrix} = \begin{bmatrix} F_{11} & F_{12} & \cdots & F_{1N} \\ \vdots & \vdots & & \vdots \\ F_{i1} & F_{i2} & \cdots & F_{iN} \\ \vdots & \vdots & & \vdots \\ F_{N1} & F_{N2} & \cdots & F_{NN} \end{bmatrix} \begin{pmatrix} y_1^o \\ y_2^o \\ \vdots \\ y_N^o \end{pmatrix}$$

Veja então que s se a matriz densidade de informação for diferente da matriz identidade, $\bar{\bar{\mathbf{F}}} \neq \bar{\bar{\mathbf{I}}}$ então o j-ésimo elemento do vetor de dados estimados, $\underset{\sim}{\tilde{\mathbf{y}}}$, é

$$\tilde{y}_i = F_{i1} y_1^o + F_{i2} y_2^o + \cdots + F_{iN} \, y_N^o$$

$$\tilde{y}_i = \sum_{j=1}^{N} F_{ij} y_j^o \; .$$