Revisão: Inversão linear

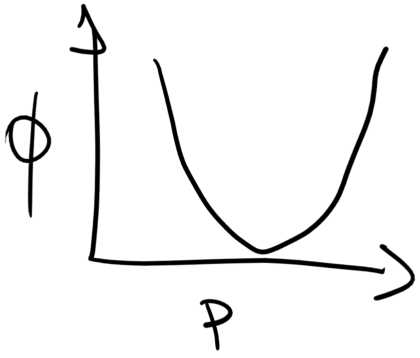
$$\bar{d} = \bar{f}(\bar{p}) = \bar{\bar{A}}\,\bar{p}$$

$\bar{\bar{A}}$ = Jacobiana

$$\begin{bmatrix} \partial_{p_1} f_1 & \partial_{p_2} f_1 & \partial_{p_3} f_1 & \cdots \\ \partial_{p_1} f_2 & \partial_{p_2} f_2 & \cdots & \\ & & \ddots & \\ & & & \ddots \end{bmatrix}$$

constante

$$\phi(\bar{p}) = \| \bar{d}^o - \bar{\bar{A}}\bar{p} \|^2 = \underbrace{[\bar{d}^o - \bar{\bar{A}}\bar{p}]^T[\bar{d}^o - \bar{\bar{A}}\bar{p}]}_{\text{forma quadrática}}$$



$$\min_{\bar{p}} \phi \Rightarrow \bar{\nabla}\phi = \bar{0}$$

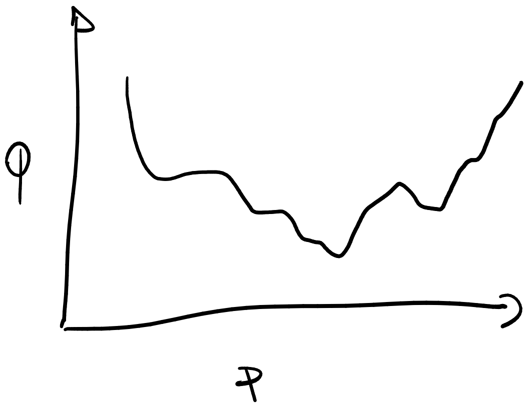
$$\bar{\nabla}\phi = \bar{\bar{A}}^T\bar{\bar{A}}\bar{p} - \bar{\bar{A}}^T\bar{d}^o$$

$$\underbrace{\bar{\bar{A}}^T\bar{\bar{A}}\bar{p} = \bar{\bar{A}}^T\bar{d}^o}_{\text{sistema de equações normais}}$$

Inversão não-linear

$$\bar{d} = \bar{f}(\bar{p})$$

$$\phi(\bar{p}) = [\bar{d}^o - \bar{f}(\bar{p})]^T [\bar{d}^o - \bar{f}(\bar{p})]$$



Método de Newton

Série de Taylor

$$\phi(\bar{p}) = \underbrace{\phi(\bar{p}_0) + \bar{\nabla}\phi(\bar{p}_0)\bar{\Delta p} + \frac{1}{2}\bar{\Delta p}^T \bar{\bar{H}} \bar{\Delta p}}_{\psi} + \dots$$

$$\phi(\bar{p}) \approx \psi(\bar{\Delta p}) \qquad \bar{\bar{H}} = \bar{\nabla}\bar{\nabla}^T\phi = \text{Hessiana}$$

$$\min_{\bar{\Delta p}} \psi \Rightarrow \bar{\nabla}_{\Delta p}\psi = \bar{0} = \bar{\nabla}_p\phi + \frac{1}{\cancel{2}}\cancel{2}\bar{\bar{H}}\bar{\Delta p}$$

$$\bar{\bar{H}}\bar{\Delta p} = -\bar{\nabla}_p\phi$$

sistema de equações normais

Processo:

1. $\bar{p}_0$
2. $\bar{p}_0 \Rightarrow \bar{\Delta p} \Rightarrow \bar{p}_1 = \bar{p}_0 + \bar{\Delta p}$
3. Repete até Niter ou convergência

$\bar{\nabla}\phi$? $\bar{\bar{A}}$?

$$\bar{\nabla}\phi = \bar{\nabla}\left[\bar{d}^{o} - \bar{f}(\bar{p})\right]^{T}\left[\bar{d}^{o} - \bar{f}(\bar{p})\right] = \bar{\nabla}\left[\bar{d}^{oT}\bar{d}^{o} + \bar{f}^{T}\bar{f} - 2\bar{f}^{T}\bar{d}^{o}\right] = \bar{\nabla}f^{T}f - 2\bar{\nabla}\bar{f}^{T}\bar{d}^{o}$$

$$\bar{\nabla}f^{T} = \begin{bmatrix} \partial p_1 \\ \vdots \end{bmatrix}\begin{bmatrix} f_1 & f_2 & \cdots \end{bmatrix} = \begin{bmatrix} \partial_{p_1} f_1 & \partial_{p_1} f_2 & \cdots \\ \partial_{p_2} f_1 & \partial_{p_2} f_2 & \cdots \\ \vdots & & \end{bmatrix} = \bar{\bar{A}}^{T}$$

$$\bar{\nabla}f^{T}f = \begin{bmatrix} \partial_{p_1} \\ \partial_{p_2} \\ \vdots \end{bmatrix}\left(f_1^2 + f_2^2 + f_3^2 \cdots\right) = \begin{bmatrix} 2f_1\partial_{p_1}f_1 + 2f_2\partial_{p_1}f_2 & \cdots \\ 2f_1\partial_{p_2}f_1 + 2f_2\partial_{p_2}f_2 & \cdots \\ \vdots & \end{bmatrix} = 2\begin{bmatrix} \partial_{p_1}\bar{f}^{T}\bar{f} \\ \partial_{p_2}\bar{f}^{T}\bar{f} \\ \vdots \end{bmatrix} = 2\underbrace{\begin{bmatrix} \partial_{p_1}\bar{f}^{T} \\ \partial_{p_2}\bar{f}^{T} \\ \vdots \end{bmatrix}}_{\bar{\bar{A}}^{T}}\bar{f}$$

$$\bar{\nabla}\phi = 2\bar{\bar{A}}^{T}\bar{f} - 2\bar{\bar{A}}^{T}\bar{d}^{o} = -2\bar{\bar{A}}^{T}(\bar{d}^{o} - \bar{f})$$

$$\bar{H} = \bar{\nabla}\,\bar{\nabla}^T\phi = \bar{\nabla}\left(-2\bar{A}^T(\bar{d}^o - \bar{f})\right)^T = -2\begin{bmatrix}\partial_{p_1}\\ \partial_{p_2}\\ \vdots\end{bmatrix}\begin{bmatrix} a_1^T(\bar{d}^o-\bar{f}) & a_2^T(\bar{d}^o-\bar{f}) & \cdots\end{bmatrix}$$

$$H_{ij} = \partial_{p_i}\,\bar{a}_j^T(\bar{d}^o - \bar{f})$$

$$= \partial_{p_i}\left(\partial_{p_j}\bar{f}^T(\bar{d}^o - \bar{f})\right)$$

$$= \cancelto{0}{\partial_{p_i}\partial_{p_j}\bar{f}^T(\bar{d}^o - \bar{f})} - \partial_{p_j}\bar{f}^T\partial_{p_i}\bar{f}$$

$$\approx -\partial_{p_j}\bar{f}^T\partial_{p_i}\bar{f} = -\partial_{p_i}\bar{f}^T\partial_{p_j}\bar{f}$$

$$\bar{\bar{H}} = 2\begin{bmatrix}\partial_{p_1}\bar{f}^T\partial_{p_1}\bar{f} & \partial_{p_1}\bar{f}^T\partial_{p_2}\bar{f} & \cdots\\ \partial_{p_2}\bar{f}^T\partial_{p_1}\bar{f} & \partial_{p_2}\bar{f}^T\partial_{p_2}\bar{f} & \cdots\\ & \ddots & \end{bmatrix}$$

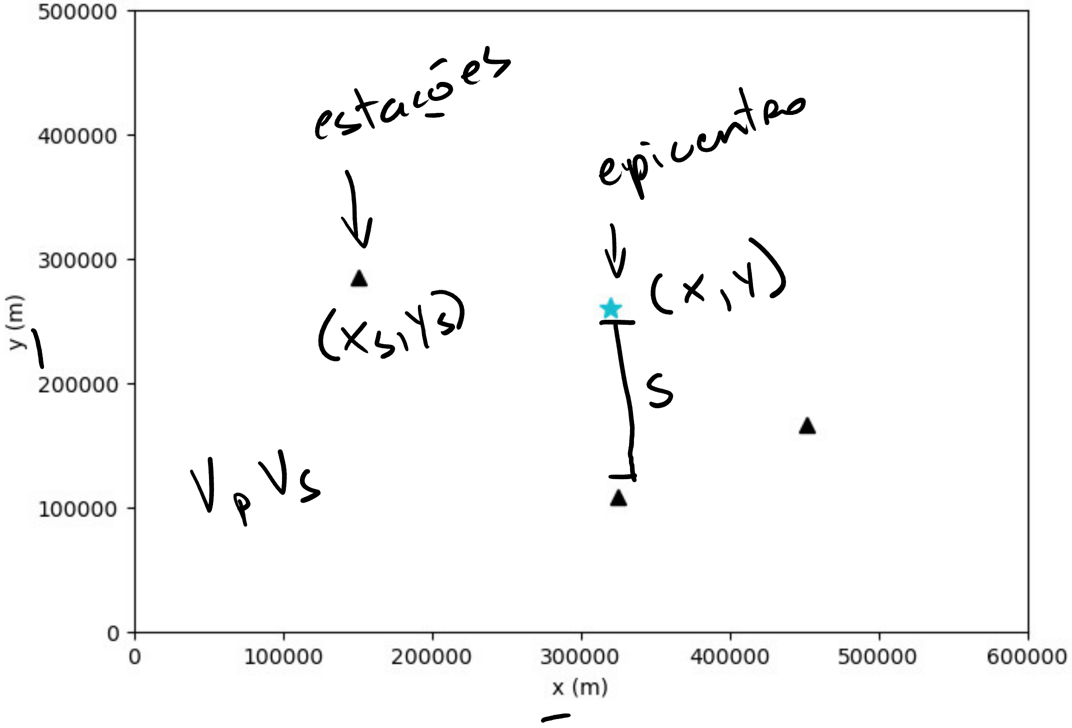
$$= 2\begin{bmatrix}\partial_{p_1}\bar{f}^T\\ \partial_{p_2}\bar{f}^T\\ \vdots\end{bmatrix}\begin{bmatrix}\partial_{p_1}\bar{f} & \partial_{p_2}\bar{f} & \cdots\end{bmatrix}$$

$$= \bar{\bar{A}}^T\bar{\bar{A}}$$

$$\bar{\bar{H}}\,\overline{\Delta p} = -\overline{\nabla}\phi \quad \Rightarrow \quad \cancel{\alpha}\,\bar{\bar{A}}^T\bar{\bar{A}}\,\overline{\Delta p} = -\left(-\cancel{\alpha}\,\bar{\bar{A}}^T(\bar{d}^0-\bar{f})\right) = \underline{\bar{\bar{A}}^T\bar{\bar{A}}\,\overline{\Delta p} = \bar{\bar{A}}^T(\bar{d}^0-\bar{f})}$$

Relação entre linear e não-linear

Exemplo: Localização de epicentros de terremotos



$$t_p = \frac{s}{V_p} \qquad t_s = \frac{s}{V_s} \qquad s = \sqrt{(x-x_s)^2 + (y-y_s)^2}$$

$$\Delta t = \sqrt{(x-x_s)^2 + (y-y_s)^2}\left(\frac{1}{V_s} - \frac{1}{V_p}\right)$$

$$\downarrow$$

$$\bar{p} = \begin{bmatrix} x \\ y \end{bmatrix} \qquad \text{não-linear} \Rightarrow \bar{\bar{A}}\ ?$$

$$A_{ij} = \frac{\partial \Delta t_i}{\partial p_j}$$

$$\frac{\partial \Delta t}{\partial x} = \frac{1}{2}\frac{1}{s}2x\left(\frac{1}{V_s} - \frac{1}{V_p}\right)$$

$$\frac{\partial \Delta t}{\partial x} = \frac{x}{s}\left(\frac{1}{V_s} - \frac{1}{V_p}\right)$$

Para o Python!

Steepest descent

Andar na direção contrária ao gradiente

$$\bar{\Delta}p = -\underline{\lambda}\,\bar{\nabla}\phi$$

$\downarrow$

passo $\Rightarrow$ determinado por "line search"

$\Downarrow$

$$\lambda = \beta^{l}$$

Levemberg-Marquardt

Controla o tamanho do passo no método de Newton

$$\underbrace{\Gamma(\bar{p})}_{} = \underbrace{\Psi(\bar{p})}_{\text{série de Taylor}} + \underbrace{\alpha \bar{\Delta p}^T \bar{\Delta p}}_{\text{linear}}$$

série de Taylor

linear

controla o tamanho de $\bar{\Delta p}$

$$\min_{\bar{\Delta p}} \Gamma \Rightarrow \bar{\nabla} \Gamma = \bar{0}$$

$$\bar{\nabla}\Gamma = \bar{\nabla}\phi + \bar{\bar{H}}\Delta p + 2\alpha \bar{\Delta p} = \bar{0}$$

$$(\cancel{2}\bar{\bar{A}}^T\bar{\bar{A}} + \cancel{2}\alpha \bar{\bar{I}})\bar{\Delta p} = -(-\cancel{2}\bar{\bar{A}}^T(\bar{d}^0 - \bar{f}))$$

$$(\bar{\bar{A}}^T\bar{\bar{A}} + \alpha \bar{\bar{I}})\bar{\Delta p} = \bar{\bar{A}}^T(\bar{d}^0 - \bar{f})$$

Procedimento prático:

1. começa com α pequeno
2. se ϕ aumenta $\alpha = \alpha \Delta\alpha$ $\Rightarrow$ maior α = menor $\Delta\bar{p}$
3. Calcule novo $\Delta\bar{p}$
4. se ϕ diminui $\bar{p}_k = \bar{p}_{k-1} + \Delta\bar{p}$ e $\alpha = \alpha / \Delta\alpha$

Conclusão

Inversão não-linear $\Rightarrow$ métodos iterativos de buscar o mínino de $\phi(\bar{p})$

Newton:
1. Aproxima função objetivo por formas quadráticas
2. Revolve vários problemas lineares
3. Problemas de convergência

Steepest descent:
1. Anda na direção oposta ao gradiente
2. Não resolve sistemas lineares
3. Mas precisa de mais passos para convergir

Levemberg-Marquardt:
1. Regularização do tamanho do passo
2. Boa convergência
3. Mas precisa resolver sistema de eq. normais várias vezes