Fitting implicit models

model $\Rightarrow$ $\bar{f}(\bar{d}, \bar{p}) = \bar{0}$

$\bar{f} = L \times 1$ relations

$\bar{d} = N \times 1$ data

$\bar{p} = M \times 1$ model parameters

Goal function

$$\Gamma(\bar{d}, \bar{p}, \bar{\lambda}) = \underbrace{\phi(\bar{d})}_{\substack{\text{data misfit} \\ \text{data constraints}}} + \underbrace{\psi(\bar{p})}_{\substack{\text{Regularization} \\ \text{model constraints}}} + \overbrace{\bar{\lambda}}^{\text{Lagrange multipliers}}{}^T \underbrace{f(\bar{d}, \bar{p})}_{\substack{\text{forward} \\ \text{model}}}$$

$$\bar{u} = \begin{bmatrix} \bar{d} \\ \bar{\lambda} \\ \bar{p} \end{bmatrix} \qquad \hat{\bar{u}} = \min \Gamma(\bar{u})$$

Newton's method for optimization

Taylor series: $\Gamma'(\bar{u}) = \Gamma(\bar{u}^0) + \underbrace{\bar{\nabla}_u \Gamma(\bar{u}^0)}_{\text{gradient}}{}^T \bar{\Delta u} + \frac{1}{2} \bar{\Delta u}^T \underbrace{\bar{\bar{\nabla}}_u \Gamma(\bar{u}^0)}_{\text{Hessian}} \bar{\Delta u} + \overbrace{O(3)}^{\text{ignore}}$

$$\min \Gamma'(\bar{u}) \Rightarrow \bar{\nabla}_u \Gamma' = \bar{0} \Rightarrow \bar{\nabla}_u \Gamma(\bar{u}^0) + \bar{\bar{\nabla}}_u \Gamma(\bar{u}^0) \bar{\Delta u} = \bar{0}$$

$$\underbrace{\bar{\bar{\nabla}}_u \Gamma(\bar{u}^0) \bar{\Delta u} = - \bar{\nabla}_u \Gamma(\bar{u}^0)}_{\text{Normal equations}}$$

Repeat iteratively until convergence

Calculating the gradient

$$\bar{\nabla}_u \Gamma = \begin{bmatrix} \bar{\nabla}_d \Gamma \\ \bar{\nabla}_\lambda \Gamma \\ \bar{\nabla}_p \Gamma \end{bmatrix}$$

$$\bar{\nabla}_\lambda \Gamma = \bar{f}(\bar{d}, \bar{p}) \Rightarrow \boxed{\bar{\nabla}_\lambda \Gamma(\bar{u}^\circ) = \bar{f}^\circ}$$

$$\bar{\nabla}_d \Gamma = \bar{\nabla}_d \phi + \cancel{\bar{\nabla}_d \psi}^{\,0} + \bar{\nabla}_d \bar{\lambda}^T \bar{f}(\bar{d}, \bar{p}) \Rightarrow \boxed{\bar{\nabla}_d \Gamma(\bar{u}^\circ) = \bar{\nabla}_d \phi(\bar{u}^\circ) + \bar{\bar{B}}(\bar{u}^\circ)^T \bar{\lambda}}$$

$$\bar{\nabla}_d \bar{\lambda}^T \bar{f} = \bar{\nabla}_d \bar{f}^T \bar{\lambda} = \begin{bmatrix} \frac{\partial \bar{f}}{\partial d_1}^T \\ \frac{\partial \bar{f}}{\partial d_2}^T \\ \vdots \\ \frac{\partial \bar{f}}{\partial d_N}^T \end{bmatrix} \bar{\lambda} = \bar{\bar{B}}^T \bar{\lambda}$$

$N \times L$

Data Jacobian

$$\bar{\nabla}_p \Gamma = \cancel{\bar{\nabla}_p \phi}^{\,0} + \bar{\nabla}_p \psi + \bar{\nabla}_p \bar{\lambda}^T \bar{f}(\bar{d}, \bar{p}) \Rightarrow \boxed{\bar{\nabla}_p \Gamma(\bar{u}^\circ) = \bar{\nabla}_p \psi(\bar{u}^\circ) + \bar{\bar{A}}(\bar{u}^\circ)^T \bar{\lambda}}$$

parameter Jacobian

Calculating the Hessian

$$\bar{\bar{\nabla}}_u \Gamma = \bar{\bar{H}}_u$$

$$\bar{\bar{H}}_u = \begin{bmatrix} \bar{\bar{H}}_d & \bar{\bar{H}}_{d\lambda} & \bar{\bar{H}}_{dp} \\ \bar{\bar{H}}_{\lambda d} & \bar{\bar{H}}_\lambda & \bar{\bar{H}}_{\lambda p} \\ \bar{\bar{H}}_{pd} & \bar{\bar{H}}_{p\lambda} & \bar{\bar{H}}_p \end{bmatrix} \Rightarrow \bar{\bar{H}}_{dp}^T = \bar{\bar{H}}_{pd} \Rightarrow \begin{bmatrix} \bar{\bar{H}}_d & \bar{\bar{H}}_{d\lambda} & \bar{\bar{H}}_{dp} \\ \bar{\bar{H}}_{d\lambda}^T & \bar{\bar{H}}_\lambda & \bar{\bar{H}}_{\lambda p} \\ \bar{\bar{H}}_{dp}^T & \bar{\bar{H}}_{\lambda p}^T & \bar{\bar{H}}_p \end{bmatrix}$$

$\bar{\bar{H}}_d \Rightarrow N \times N$ $\quad$ $\bar{\bar{H}}_{d\lambda} \Rightarrow N \times L$

$\bar{\bar{H}}_p \Rightarrow M \times M$ $\quad$ $\bar{\bar{H}}_{dp} \Rightarrow N \times M$

$\bar{\bar{H}}_\lambda \Rightarrow L \times L$ $\quad$ $\bar{\bar{H}}_{\lambda p} \Rightarrow L \times M$

$\bar{\bar{H}}_\lambda = \bar{\bar{0}}$ since Γ is linear with $\bar{\lambda}$

$$\bar{\bar{H}}_{d\lambda} = \bar{\nabla}_d\left[\bar{\nabla}_\lambda \Gamma^T\right] = \bar{\nabla}_d \bar{f}^T = \begin{bmatrix} \frac{\partial \bar{f}^T}{\partial d_1} \\ \vdots \\ \frac{\partial \bar{f}^T}{\partial d_N} \end{bmatrix} = \bar{\bar{B}}^T$$

$$\bar{\bar{H}}_{p\lambda} = \bar{\nabla}_p\left[\bar{\nabla}_\lambda \Gamma^T\right] = \bar{\nabla}_p \bar{f}^T = \bar{\bar{A}}^T$$

$$\bar{\bar{H}}_{dp} = \bar{\nabla}_d\left[\bar{\nabla}_p \Gamma^T\right] = \cancelto{0}{\bar{\nabla}_d \bar{\nabla}_p \psi^T} + \underbrace{\bar{\nabla}_d\left[\bar{\bar{A}}^T \bar{\lambda}\right]^T}_{\substack{2^{nd} \text{ derivatives} \\ \text{of } \bar{f} \Rightarrow 0 \\ \text{Gauss-Newton}}} \approx \bar{\bar{0}}$$

$$\bar{\bar{H}}_{d} = \bar{\nabla}_d\left[\bar{\nabla}_d \Gamma^T\right] = \bar{\nabla}_d\left[\bar{\nabla}_d \phi^T\right] + \cancelto{0}{\bar{\nabla}_d\left[\bar{\bar{B}}^T \bar{\lambda}\right]^T} \approx \bar{\nabla}_d \phi$$

$$\bar{\bar{H}}_{p} = \bar{\nabla}_p\left[\bar{\nabla}_p \Gamma^T\right] = \bar{\nabla}_p\left[\bar{\nabla}_p \psi^T\right] + \cancelto{0}{\bar{\nabla}_p\left[\bar{\bar{A}}^T \bar{\lambda}\right]^T} \approx \bar{\nabla}_p \psi$$

Normal equations

System 1

$$\left[\begin{array}{c|cc} \bar{\bar{H}}_d & \bar{\bar{B}}^T & \bar{\bar{0}} \\ \hline \bar{\bar{B}} & \bar{\bar{0}} & \bar{\bar{A}} \\ \bar{\bar{0}} & \bar{\bar{A}}^T & \bar{\bar{H}}_p \end{array}\right] \left[\begin{array}{c} \bar{\Delta d} \\ \hline \bar{\Delta \lambda} \\ \bar{\Delta p} \end{array}\right] + \left[\begin{array}{c} \bar{\nabla}_d \phi + \bar{\bar{B}}^T \bar{\lambda} \\ \hline \bar{f} \\ \bar{\nabla}_p \psi + \bar{\bar{A}}^T \lambda \end{array}\right] = \left[\begin{array}{c} \bar{0} \\ \hline \bar{0} \\ \bar{0} \end{array}\right]$$

$$\hookrightarrow \begin{bmatrix} \bar{\bar{A}} & \bar{\bar{B}} \\ \bar{\bar{C}} & \bar{\bar{D}} \end{bmatrix} \begin{bmatrix} \bar{x} \\ \bar{y} \end{bmatrix} + \begin{bmatrix} \bar{u} \\ \bar{v} \end{bmatrix} = \begin{bmatrix} \bar{0} \\ \bar{0} \end{bmatrix} \Rightarrow \left[\bar{\bar{D}} - \bar{\bar{C}}\bar{\bar{A}}^{-1}\bar{\bar{B}}\right]\bar{y} + \bar{v} - \bar{\bar{C}}\bar{\bar{A}}^{-1}\bar{u} = \bar{0}$$

$$\left[\begin{bmatrix}\bar{\bar{0}} & \bar{\bar{A}} \\ \bar{\bar{A}}^T & \bar{\bar{H}}_p\end{bmatrix} - \begin{bmatrix}\bar{\bar{B}} \\ \bar{\bar{0}}\end{bmatrix}\bar{\bar{H}}_d^{-1}\begin{bmatrix}\bar{\bar{B}}^T & \bar{\bar{0}}\end{bmatrix}\right]\begin{bmatrix}\bar{\Delta\lambda} \\ \bar{\Delta p}\end{bmatrix} + \begin{bmatrix}\bar{f} \\ \bar{\nabla}_p\psi + \bar{\bar{A}}^T\bar{\lambda}\end{bmatrix} - \begin{bmatrix}\bar{\bar{B}} \\ \bar{\bar{0}}\end{bmatrix}\bar{\bar{H}}_d^{-1}\left(\bar{\nabla}_d\phi + \bar{\bar{B}}^T\bar{\lambda}\right) = \begin{bmatrix}\bar{0} \\ \bar{0}\end{bmatrix}$$

system 2

$$\left[\begin{array}{c|c} -\bar{\bar{B}}\bar{\bar{H}}_d^{-1}\bar{\bar{B}}^T & \bar{\bar{A}} \\ \hline \bar{\bar{A}}^T & \bar{\bar{H}}_p \end{array}\right]\left[\begin{array}{c}\Delta\bar{\lambda} \\ \hline \bar{\Delta p}\end{array}\right] + \left[\begin{array}{c}\bar{f} - \bar{\bar{B}}\bar{\bar{H}}_d^{-1}\bar{\nabla}_d\phi - \bar{\bar{B}}\bar{\bar{H}}_d^{-1}\bar{\bar{B}}^T\bar{\lambda} \\ \hline \bar{\nabla}_p\psi + \bar{\bar{A}}^T\bar{\lambda}\end{array}\right] = \left[\begin{array}{c}\bar{0} \\ \hline \bar{0}\end{array}\right]$$

$$\left[\bar{\bar{H}}_p + \bar{\bar{A}}^T\left[\bar{\bar{B}}\bar{\bar{H}}_d^{-1}\bar{\bar{B}}^T\right]^{-1}\bar{\bar{A}}\right]\Delta\bar{p} + \bar{\nabla}_p\psi + \bar{\bar{A}}^T\bar{\lambda} + \bar{\bar{A}}^T\left[\bar{\bar{B}}\bar{\bar{H}}_d^{-1}\bar{\bar{B}}^T\right]^{-1}\left[\bar{f} - \bar{\bar{B}}\bar{\bar{H}}_d^{-1}\bar{\nabla}_d\phi - \bar{\bar{B}}\bar{\bar{H}}_d^{-1}\bar{\bar{B}}^T\bar{\lambda}\right] = \bar{0}$$

$$\left[\bar{\bar{H}}_p + \bar{\bar{A}}^T\left[\bar{\bar{B}}\bar{\bar{H}}_d^{-1}\bar{\bar{B}}^T\right]^{-1}\bar{\bar{A}}\right]\Delta\bar{p} + \bar{\nabla}_p\psi + \cancel{\bar{\bar{A}}^T\lambda} + \bar{\bar{A}}^T\left[\bar{\bar{B}}\bar{\bar{H}}_d^{-1}\bar{\bar{B}}^T\right]^{-1}\left[\bar{f} - \bar{\bar{B}}\bar{\bar{H}}_d^{-1}\bar{\nabla}_d\phi\right] - \cancel{\bar{\bar{A}}^T\lambda} = \bar{0}$$

$$\boxed{\left[\bar{\bar{H}}_p + \bar{\bar{A}}^T\left[\bar{\bar{B}}\bar{\bar{H}}_d^{-1}\bar{\bar{B}}^T\right]^{-1}\bar{\bar{A}}\right]\Delta\bar{p} = \bar{\bar{A}}^T\left[\bar{\bar{B}}\bar{\bar{H}}_d^{-1}\bar{\bar{B}}^T\right]^{-1}\left[\bar{\bar{B}}\bar{\bar{H}}_d^{-1}\bar{\nabla}_d\phi - \bar{f}\right] - \bar{\nabla}_p\psi}$$

Solution for $\overline{\Delta p}$

Go back to system 2 to solve for $\overline{\Delta\lambda}$

$$-\bar{\bar{B}}\bar{\bar{H}}_d^{-1}\bar{\bar{B}}^T\Delta\lambda + \bar{\bar{A}}\Delta p + \bar{f} - \bar{\bar{B}}\bar{\bar{H}}_d^{-1}\bar{\nabla}_d\phi - \bar{\bar{B}}\bar{\bar{H}}_d^{-1}\bar{\bar{B}}^T\bar{\lambda} = \bar{0}$$

$$\boxed{\overline{\Delta\lambda} = \left[\bar{\bar{B}}\bar{\bar{H}}_d^{-1}\bar{\bar{B}}^T\right]^{-1}\left[\bar{\bar{A}}\Delta\bar{p} + \bar{f} - \bar{\bar{B}}\bar{\bar{H}}_d^{-1}\bar{\nabla}_d\phi\right] - \bar{\lambda}}$$

Go back to system 1 to solve for Δd

$$\bar{\bar{H}}_d\Delta\bar{d} + \bar{\bar{B}}^T\overline{\Delta\lambda} + \bar{\bar{0}}\,\Delta\bar{p} + \bar{\nabla}_d\phi + \bar{\bar{B}}^T\bar{\lambda} = \bar{0}$$

$$\bar{\bar{H}}_d\Delta\bar{d} + \bar{\bar{B}}^T\left[\bar{\bar{B}}\bar{\bar{H}}_d^{-1}\bar{\bar{B}}^T\right]^{-1}\left[\bar{\bar{A}}\Delta\bar{p} + \bar{f} - \bar{\bar{B}}\bar{\bar{H}}_d^{-1}\bar{\nabla}_d\phi\right] - \cancel{\bar{\bar{B}}^T}\bar{\lambda} + \bar{\nabla}_d\phi + \bar{\bar{B}}^T\cancel{\bar{\lambda}} = \bar{0}$$

$$\bar{\bar{H}}_d\Delta\bar{d} + \bar{\bar{B}}^T\left[\bar{\bar{B}}\bar{\bar{H}}_d\bar{\bar{B}}^T\right]^{-1}\left[\bar{\bar{A}}\Delta\bar{p} + \bar{f}\right] + \left[\bar{\bar{I}} - \bar{\bar{B}}^T\left[\bar{\bar{B}}\bar{\bar{H}}_d^{-1}\bar{\bar{B}}^T\right]^{-1}\bar{\bar{B}}\bar{\bar{H}}_d^{-1}\right]\bar{\nabla}_d\phi = \bar{0}$$

$$\boxed{\Delta\bar{d} = \bar{\bar{H}}_d^{-1}\left[\left[\bar{\bar{B}}^T\left[\bar{\bar{B}}\bar{\bar{H}}_d^{-1}\bar{\bar{B}}^T\right]^{-1}\bar{\bar{B}}\bar{\bar{H}}_d^{-1} - \bar{\bar{I}}\right]\bar{\nabla}_d\phi - \bar{\bar{B}}^T\left[\bar{\bar{B}}\bar{\bar{H}}_d\bar{\bar{B}}^T\right]^{-1}\left[\bar{\bar{A}}\Delta\bar{p} + \bar{f}\right]\right]}$$

Regulating the step size $\Delta\bar{u}$

$$\Gamma'(\bar{u}) = \Gamma(\bar{u}_0) + \bar{\nabla}_u \Gamma(\bar{u}_0)\,\Delta\bar{u} + \frac{1}{2}\Delta\bar{u}^T\,\bar{\bar{\nabla}}_u\Gamma(\bar{u}_0)\,\Delta\bar{u} + \overbrace{\frac{\alpha}{2}\Delta\bar{u}^T\,\bar{\bar{W}}_u\,\Delta\bar{u}}^{\text{regularization for } \Delta\bar{u}}$$

$$\bar{\nabla}_u\Gamma' = \bar{\nabla}\Gamma(\bar{u}_0) + \bar{\bar{\nabla}}_u\Gamma(\bar{u}_0)\,\Delta u + \alpha\,\bar{\bar{W}}_u\,\Delta\bar{u} = \bar{\nabla}_u\Gamma(\bar{u}_0) + \left[\bar{\bar{\nabla}}_u\Gamma(\bar{u}_0) + \alpha\,\bar{\bar{W}}_u\right]\Delta\bar{u} = \bar{0}$$

Normal equations

$$\left[\begin{bmatrix} \bar{\bar{H}}_d & \bar{\bar{B}}^T & \bar{\bar{0}} \\ \bar{\bar{B}} & \bar{\bar{0}} & \bar{\bar{A}} \\ \bar{\bar{0}} & \bar{\bar{A}}^T & \bar{\bar{H}}_p \end{bmatrix} + \begin{bmatrix} \alpha\bar{\bar{W}}_d & \bar{\bar{0}} & \bar{\bar{0}} \\ \bar{\bar{0}} & \bar{\bar{0}} & \bar{\bar{0}} \\ \bar{\bar{0}} & \bar{\bar{0}} & \alpha\bar{\bar{W}}_p \end{bmatrix}\right] \begin{bmatrix} \Delta\bar{d} \\ \Delta\bar{\lambda} \\ \Delta\bar{p} \end{bmatrix} + \begin{bmatrix} \bar{\nabla}_d\phi + \bar{\bar{B}}^T\bar{\lambda} \\ \bar{f} \\ \bar{\nabla}_p\psi + \bar{\bar{A}}^T\bar{\lambda} \end{bmatrix} = \begin{bmatrix} \bar{0} \\ \bar{0} \\ \bar{0} \end{bmatrix}$$

$$\left[\bar{\bar{A}}^T\left[\bar{\bar{B}}\left[\bar{\bar{H}}_d + \alpha\bar{\bar{W}}_d\right]^{-1}\bar{\bar{B}}^T\right]^{-1}\bar{\bar{A}} + \bar{\bar{H}}_p + \alpha\bar{\bar{W}}_p\right]\Delta p = \bar{\bar{A}}^T\left[\bar{\bar{B}}\left[\bar{\bar{H}}_d + \alpha\bar{\bar{W}}_d\right]^{-1}\bar{\bar{B}}^T\right]^{-1}\left[\bar{\bar{B}}\left[\bar{\bar{H}}_d + \alpha\bar{\bar{W}}_d\right]^{-1}\bar{\nabla}_d\phi - \bar{f}\right] - \bar{\nabla}_p\psi$$

$$\left[\bar{\bar{B}}\left[\bar{\bar{H}}_d + \alpha\bar{\bar{W}}_d\right]^{-1}\bar{\bar{B}}^T\right]^{-1} = \bar{\bar{Q}} \;\Rightarrow\; \underline{\left[\bar{\bar{A}}^T\bar{\bar{Q}}\bar{\bar{A}} + \bar{\bar{H}}_p + \alpha\bar{\bar{W}}_p\right]\Delta\bar{p} = \bar{\bar{A}}^T\bar{\bar{Q}}\left[\bar{\bar{B}}\left[\bar{\bar{H}}_d + \alpha\bar{\bar{W}}_d\right]^{-1}\bar{\nabla}_d\phi - \bar{f}\right] - \bar{\nabla}_p\psi}$$

$$\underline{\Delta\bar{d} = \left[\bar{\bar{H}}_d + \alpha\bar{\bar{W}}_d\right]^{-1}\left[\left[\bar{\bar{B}}^T\bar{\bar{Q}}\,\bar{\bar{B}}\left[\bar{\bar{H}}_d + \alpha\bar{\bar{W}}_d\right]^{-1} - \bar{\bar{I}}\right]\bar{\nabla}_d\phi - \bar{\bar{B}}^T\bar{\bar{Q}}\left[\bar{\bar{A}}\,\Delta\bar{p} + \bar{f}\right]\right]}$$