FOR AUGUST 2022 WORKSHOP

1. Linearization

Expanding to first order about $\mathbf{x}$:

$$\mathbf{f}(\mathbf{x}+\mathbf{\Delta x}) \approx \mathbf{f}(\mathbf{x}) + \mathbf{J}\mathbf{\Delta x}, \tag{1}$$

where Jacobian $\mathbf{J} \in \mathbb{R}^{n_f \times n_x}$. Now writing residual $\mathbf{r} = \mathbf{f}(\mathbf{x}+\mathbf{\Delta x}) - \mathbf{f}(\mathbf{x}) \approx \mathbf{d} - \mathbf{f}(\mathbf{x})$,

$$\mathbf{J}\mathbf{\Delta x} = \mathbf{r}, \tag{2}$$

which is a linear least squares problem with solution:

$$\mathbf{\Delta x} = (\mathbf{J}^t\mathbf{J})^{-1}\mathbf{J}^t\mathbf{r} \tag{3}$$

For an objective function $\phi = ||\mathbf{d} - \mathbf{f}(\mathbf{x})||$, you can prove with a little multivariable calculus that gradient $\nabla_x\phi = -\mathbf{J}^t\mathbf{r}$, which sort of all begins to make sense now, as the new

$$\mathbf{x}' = \mathbf{x} - \gamma\nabla_x\phi \tag{4}$$

$$\mathbf{x}' = \mathbf{x} + \mathbf{\Delta x} \tag{5}$$

with Hessian inverse approximated by $(\mathbf{J}^t\mathbf{J})^{-1}$ as step length γ. The direction to walk down (since gradient points up) can be identified with $-\nabla_x\phi$ given by $\mathbf{J}^t\mathbf{r}$. If we are not in the vicinity of the minimum we repeat this step various times, with a new linearization about $\mathbf{x}'$.

2. A system

$$\begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ \vdots \\ y_n \end{bmatrix} = \begin{bmatrix} 1 & 0 & \cdots & 0 & 0 \\ 0 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & 0 \\ 0 & 0 & \cdots & 0 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ \vdots \\ x_n \end{bmatrix} \tag{6}$$

$$\begin{bmatrix} y_2 \end{bmatrix} = \begin{bmatrix} 0 & 1 & \cdots & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ \vdots \\ x_n \end{bmatrix} = [\,\mathbf{e}_2\,] \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ \vdots \\ x_n \end{bmatrix} \tag{7}$$

$$\begin{bmatrix} y_2 \\ y_4 \\ y_9 \end{bmatrix} = \begin{bmatrix} \mathbf{e}_2 \\ \mathbf{e}_4 \\ \mathbf{e}_9 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ \vdots \\ x_n \end{bmatrix} \tag{8}$$

$$\phi = ||\mathbf{y} - \mathbf{A}\mathbf{x}||^2, \tag{9}$$

$$\text{set } \nabla_x\phi = 0, \tag{10}$$

$$\hat{\mathbf{x}} = (\mathbf{A}^\mathbf{t}\mathbf{A})^{-1}\mathbf{A}^\mathbf{t}\mathbf{d}. \tag{11}$$

$$\hat{\mathbf{x}}_{\text{ridge}} = (\mathbf{A}^\mathbf{t}\mathbf{A} + \delta^2\mathbf{I})^{-1}\mathbf{A}^\mathbf{t}\mathbf{d}. \tag{12}$$

$$\begin{bmatrix} & 0 & & & \\ -1 & 1 & & 0 & \\ & -1 & 1 & & \\ & & & \cdots & \\ & 0 & & -1 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} 0 \\ -x_1 + x_2 \\ -x_2 + x_3 \\ \vdots \\ -x_{n-1} x_n \end{bmatrix} \tag{13}$$

$$\phi = ||\mathbf{y} - \mathbf{A}\mathbf{x}||^2 + \lambda^2||\mathbf{R}\mathbf{x}||^2, \tag{14}$$

$$\text{set } \nabla_x \phi = 0, \tag{15}$$

$$\hat{\mathbf{x}}_{\text{smooth}} = (\mathbf{A^t A} + \lambda^\mathbf{2}\mathbf{R^t R})^{-1}\mathbf{A^t d} \tag{16}$$

$$\hat{\mathbf{x}}_{\text{occam}} = (\mathbf{A^t A} + \lambda^\mathbf{2}\mathbf{R^t R})^{-1}\mathbf{A^t d}. \tag{17}$$

$$k_i^2 = k_r^2 + k_{i_z}^2, \tag{18}$$

$$= \hat{\epsilon}_i \mu \omega^2, \tag{19}$$

$$\text{where } \hat{\epsilon}_i = \epsilon_i + i\sigma_i/\omega. \text{ But for most geophysics ...} \tag{20}$$

$$k_i^2 = \epsilon_i \mu \omega^2 + i\omega\mu\sigma_i \tag{21}$$

$$\approx i\omega\mu\sigma_i \tag{22}$$

$$\mathbf{d} = \mathbf{f}(\mathbf{m}) \tag{23}$$

$$\phi(\mathbf{m}) = \frac{1}{2}\Big(||\mathbf{d} - \mathbf{f}(\mathbf{m})||^2 + \lambda^2||\mathbf{Rm}||^2\Big), \tag{24}$$

$$\text{but how to set } \nabla_m \phi = 0\ ? \tag{25}$$

$$\textit{linearize } \phi(\mathbf{m}) \text{ to } \phi(\mathbf{m} + \mathbf{\Delta m}) \text{ i.e.,} \tag{26}$$

$$\mathbf{f}(\mathbf{m}) \to \mathbf{f}(\mathbf{m} + \mathbf{\Delta m}), \mathbf{Rm} \to \mathbf{R}(\mathbf{m} + \mathbf{\Delta m}) \text{ first.} \tag{27}$$

$$\boxed{\mathbf{f}(\mathbf{m} + \mathbf{\Delta m}) \approx \mathbf{f}(\mathbf{m}) + \mathbf{J\Delta m}.} \tag{28}$$

$$\text{first write residual } \mathbf{r} \approx \mathbf{f}(\mathbf{m}) - \mathbf{d} \tag{29}$$

$$\text{derive with respect to } \Delta\mathbf{m}, \tag{30}$$

$$\text{set } \frac{\partial \phi}{\partial \mathbf{\Delta m}} = 0, \text{ giving,} \tag{31}$$

$$\boxed{\mathbf{\Delta m} = -\Big(\mathbf{J}^t\mathbf{J} + \lambda^2\mathbf{R}^t\mathbf{R}\Big)^{-1}\Big(\mathbf{J}^t\mathbf{r} + \lambda^2\mathbf{R}^t\mathbf{Rm}\Big)} \tag{32}$$

$$\text{note also, that } \nabla_m\phi = \mathbf{J}^t\mathbf{r} + \lambda^2\mathbf{R}^t\mathbf{Rm}. \tag{33}$$

$$\text{note finally, that } \frac{\partial(\nabla_m\phi)}{\partial\mathbf{m}} = \mathbf{J}^t\mathbf{J} + \lambda^2\mathbf{R}^t\mathbf{R}. \tag{34}$$

$$\boxed{\mathbf{m}_{\text{new}} = \mathbf{m} + \mathbf{\Delta m}} \tag{35}$$

$$\text{writing } \nabla_m\phi = \mathbf{J}^t\mathbf{r} + \lambda^2\mathbf{R}^t\mathbf{Rm}, \tag{36}$$

$$\text{and } \eta = \Big(\mathbf{J}^t\mathbf{J} + \lambda^2\mathbf{R}^t\mathbf{R}\Big)^{-1} \text{ we now say,} \tag{37}$$

$$\boxed{\mathbf{m}_{\text{new}} = \mathbf{m} - \eta\nabla_m\phi} \tag{38}$$