EOSC 550: NONLINEAR INVERSION ASSIGNMENT

Consider the forward model

$$d_i = \sum_{j=1}^{N} k_j \exp(-t_i/\tau_j)$$

Download the synthetic dataset **nonlin1.txt** from the course website. The file contains synthetic data (d^{obs}), generated at twenty values of the inputs t for $N = 2$ decaying exponentials. The input file has t in the first column and d^{obs} in the second column. The data in this case have no noise.

(a) Analytically compute the Hessian H and gradient g of the misfit

$$\phi_d = \|W_d(d^{obs} - d^{pred})\|^2$$

with respect to the model parameters $m = [k_1,\ \tau_1,\ k_2,\ \tau_2]^T$. Plot H and g evaluated at the starting model $m = [1,\ 1,\ 1,\ 1]^T$. Since there is no noise in this example, we can set $W_d = I$.

(b) Newton's method: iteratively solve

$$H_k \delta m = -g_k.$$

for the model perturbation δm. Initialize with model $m = [1,\ 1,\ 1,\ 1]^T$. At each step, generate a line search which finds the minimum of $\phi_d(m_k + \alpha \delta m_k)$ with $\alpha \in [0, 1]$. You may need to monitor the condition of H to guard against small or negative eigenvalues; add a term to the diagonal $H = H + \mu I$ if necessary. The parameter μ can set as

$$\mu = abs(min(eig(H))) + percent * max(abs(eig(H)))$$

with e.g. $percent = 0.01$. Continue iterations and terminate based upon a convergence criterion (e.g. $\|\delta m\| < \epsilon$, $\|g\| < \epsilon$, $\phi_d < \epsilon$).

(c) Plot ϕ_d, α, and $\|g\|$ versus iteration. Also provide your final model and plot the observed data and the data predicted by the starting and final models. Describe the convergence criterion that you used to terminate the inversion.

(d) Repeat the minimization using the Gauss-Newton method, setting $H = J^T W_d^T W_d^T J$ (with $J_{ij} = \partial d_i / \partial m_j$, the Jacobian matrix). In this case the eigenvalues of H are guaranteed to be positive, but small eigenvalues can still occur. As in (b), if H is ill-conditioned, add a term to the diagonal $H = H + \mu I$. Report your results as in (c).

(e) Nonlinearity and ill-conditioning of the inverse problem can depend upon the definition of model parameters. For instance, we could choose $c_j = 1/\tau_j$ so that

$$d_i = \sum_{j=1}^{N} k_j \exp(-c_j t_i)$$

Minimize for $m = [k_1,\ c_1,\ k_2,\ c_2]^T$ using the Gauss-Newton algorithm. How does the choice of model parameters affect the stability of convergence?

(f) Generate a synthetic data set contaminated with Gaussian noise.

$$d_j^{obs} = d_j^{true} + \eta_j$$

The true data should be generated from your model recovered in (b). The noise should be normally distributed with zero mean and standard deviation computed as a percentage plus a floor of the true datum

$$\sigma_j = percent\ d_j^{true} + floor.$$

Use $percent = 0.05$ and $floor = 0.1$. Estimate the model using your Gauss-Newton algorithm from (e) and starting model $m = [1,\ 1,\ 1,\ 1]^T$, and report your results as in part (c). You can use the true standard deviations to define the weightings $W_{d(ii)} = 1/\sigma_i$.