6.7 Nonlinear Least Squares

$$\varphi(x) = \frac{1}{2} r^T(x) r(x) \qquad r(x) = \underset{\text{Data}}{y} - f(\underset{\text{Parameters}}{x})$$

$$\varphi(x) = \frac{1}{2} \sum_{j=1}^{n} r_j^2(x)$$

$$\frac{\partial}{\partial x_i}\left(\varphi(x)\right) = \sum_{j=1}^{n} r_j(x) \underbrace{\frac{\partial}{\partial x_i} r_j(x)}_{\left[J_r(x)\right]_{ji}}$$

$$= \left[J_r^T(x) r(x)\right]_i$$

$$H_\varphi = J(x)^T J(x) + \sum_{i=1}^{n} r_i(x) H_{r_i}(x)$$

headache

forget about it

x_k $\qquad H_\varphi(x_k)\, s = -\nabla\varphi(x_k)$

$x_{k+1} = x_k + s$

"Gauss-Newton"

$$J^T(x) J(x)\, s = -\nabla\varphi = -J^T(x) r(x)$$

$$\rightsquigarrow \quad J(x)\, s \cong -r(x)$$

$\rightarrow$ use QR to solve this least-sq. problem.

WS22p1

<u>Levenberg-Marquardt</u>

$$\left(J^T(x_k) J(x_k) + \mu_k I \right) s_k = -J^T(x_k) r(x_k)$$

equivalent LSQ problem:

$$\begin{pmatrix} J(x) \\ \sqrt{\mu} I \end{pmatrix} s \approx \begin{pmatrix} -r(x) \\ 0 \end{pmatrix}$$

$f(x,y) = \alpha x^2 + \beta y^2$

$H_f = \begin{pmatrix} 2\alpha & \\ & 2\beta \end{pmatrix}$

"Regularization"

6.8 Constrained Optimization

$$\min_x f(x) \quad \text{subject to} \quad g(x) = 0$$
$$\cancel{h(x) \leq 0}$$

"equality-constrained optimization"

Necessary cond. for min: $\nabla f(x^*) = 0$

s is a feasible direction at x:

$x + \alpha s$ is feasible for $\alpha \in [0, r]$ for some $r > 0$

interior of feasible region: every direction is feasible

Necessary condition for minimum:

$$\nabla f(x^*) \cdot s \geq 0$$

interior: s is feasible $\nabla f(x) \cdot s \geq 0$, $-s$ is feasible $\nabla f(x) \cdot (-s) \geq 0$ $\Big\}$ $\nabla f(x) = 0$

Necessary at boundary

$$-\nabla f(x^*) \in \text{row-span } J_g(x^*)$$

$$\Leftrightarrow -\nabla f(x) = J_g^T \lambda \quad \text{for some } \lambda$$

"all descent directions at our purported minimum x^* would have to be such that we'd violate the constraint, i.e. change the value of $g(x)$, if we moved in that direction"

<u>Lagrangian function</u> $\mathcal{L}(x, \lambda) = f(x) + \lambda^T g(x)$

$\nabla \mathcal{L} = 0$