Ch 10: Least-squares problems

$$f(x) = \frac{1}{2}\sum_{j=1}^{m} r_j^2(x) \quad \sim \circledast$$

$$r_j : \mathbb{R}^n \longrightarrow \mathbb{R} \,, \quad m \geq n$$

$r : \mathbb{R}^n \longrightarrow \mathbb{R}^m$:

$$r(x) = (r_1(x), r_2(x), \ldots, r_m(x))^T$$

Rewrite $\circledast$ using:

$$f(x) = \frac{1}{2}\|r(x)\|_2^2$$

$$\Rightarrow \quad J(x) = \begin{bmatrix} \nabla r_1(x)^T \\ \nabla r_2(x)^T \\ \vdots \\ \nabla r_m(x)^T \end{bmatrix}$$

"the Jacobian of r"

Also note:

$$\nabla f(x) = \sum_{j=1}^{m} r_j(x)\, \nabla r_j(x) = J(x)^T \cdot r(x)$$

$$\nabla^2 f(x) = \sum_{j=1}^{m} \nabla r_j(x)\, \nabla r_j(x)^T + \sum_{j=1}^{m} r_j(x)\, \nabla^2 r_j(x)$$

$$= J(x)^T J(x) + \sum_{j=1}^{m} r_j \nabla^2 r_j(x)$$

Basic idea:

Use $\begin{cases} \nabla f(x) = J(x)^T \cdot r(x) \\ \nabla^2 f(x) \triangleq J(x)^T J(x) \end{cases}$

10.1 Background

Work with residual: $r_j(x) = \phi(x; t_j) - y_j$

where we are trying to fit y_j.

10.2 Linear least-squares problems

$r(x) = Jx - y$ for some J, data y.

$f(x) = \frac{1}{2}\|Jx - y\|^2$, $y = r(0)$

$\nabla f(x) = J^T(Jx - y)$, $\nabla^2 f(x) = J^T J$

since $\nabla^2 r_j = 0$.

From $\nabla f(x) = 0$

$\Rightarrow$ $\boxed{J^T J\, x^* = J^T y \quad \text{normal eqns}}$

$m \geq n$ and J has full column rank.

Concluding remarks:

* Excellent algorithms do exist.
* Levenberg-Marquadt method works excellently, even for highly non-linear problems.
* Hybrid methods work for non-linear + large problems