Nonlinear Least Squares Problem

- From the assumed mathematical model $y = h(z; x_1, \ldots, x_n)$ and the observed data $\{(z_i, y_i)\}$, $i = 1, \ldots, m$, we intend to minimize the overal residual
$$R(x_1, \ldots, x_n) := \sum_{i=1}^{m} \|r_i\|_2^2,$$
where
$$r_i = r_i(x_1, \ldots, x_n) := y_i - h(z_i; x_1, \ldots, x_n).$$
- The notion can be rewritten as an unconstrained optimization problem
$$\min_{x \in R^n} F(x)$$
where
$$F(x) := \frac{1}{2}\|f(x)\|_2^2 \tag{0.1}$$
and $f(x) := [r_1(x), \ldots, r_m(x)]^T$.
- The necessary condition for $\tilde{x}$ to be a critical point is that the gradient g of F at $\tilde{x}$ is zero.
  - ⋄ We calculate the gradient of F to be
$$g(x) := \nabla F(x) = J(x)^T f(x) \tag{0.2}$$
where
$$J(x) := \frac{\partial f}{\partial x} := \begin{bmatrix} \frac{\partial r_1}{\partial x_1} & \frac{\partial r_1}{\partial x_2} & \cdots & \frac{\partial r_1}{\partial x_n} \\ \vdots & & & \\ \frac{\partial r_m}{\partial x_1} & \frac{\partial r_m}{\partial x_2} & \cdots & \frac{\partial r_m}{\partial x_n} \end{bmatrix} \tag{0.3}$$
is the $m \times n$ Jacobian matrix of f.
  - ⋄ Note that $g : R^n \rightarrow R^n$. So we may apply the Newton-Ralphson method to solve the equation $g(x) = 0$.
- Special techniques are available for this type of problems. See **leastsq** in Matlab.