5. Nonlinear least squares

- nonlinear least-squares problem
- Gauss Newton method

Nonlinear least squares

Nonlinear least squares

- The NLLS (nonlinear least-squares) problem:

$$\underset{x \in \mathbf{R}^n}{\text{minimize}} \quad \frac{1}{2}\|r(x)\|_2^2, \quad r : \mathbf{R}^n \to \mathbf{R}^m \quad \text{(typically, } m > n\text{).}$$

- "Residual" vector

$$r(x) = \begin{bmatrix} r_1(x) \\ r_2(x) \\ \vdots \\ r_m(x) \end{bmatrix}, \quad r_i : \mathbf{R}^n \to \mathbf{R}$$

differentiable functions

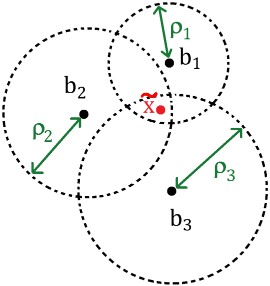
- Reduces to least-squares when r is **affine:**

$$r_i(x) = a_i^T x - b_i$$

$$r(x) = Ax - b$$

Example: position estimation from ranges

- Estimate $x \in \mathbf{R}^2$ from approximate distances to fixed beacons



- data: beacon positions

$$b_1, b_2, ..., b_m \in \mathbb{R}^2$$

- measurements

$$\rho_i = \|\tilde{x} - b_i\|_2 + v_i$$

- measurement error:

$$v_1, ..., v_m$$

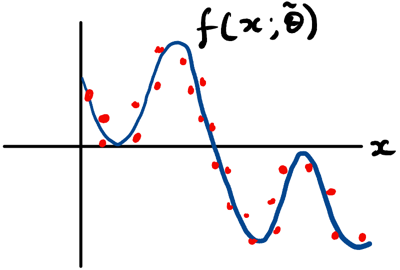
- NLLS position estimate solves

$$\underset{x \in \mathbf{R}^n}{\text{minimize}} \quad \frac{1}{2} \sum_{i=1}^{m} r_i^2(x) = \frac{1}{2} \sum_{i=1}^{m} (\rho_i - \|x - b_i\|_2)^2$$

- Must settle for locally optimal solution.

Model fitting.

Estimate model parameters $\tilde{\Theta}$ from noisy observation:



data: sampling locations

$x_1, x_2, \ldots, x_m$

non-linear model $f(x;\tilde{\Theta})$

measurement:

$$\rho_i = f(x_i;\tilde{\Theta}) + \nu_i$$

NLLS model fitting solves:

$$\min_{\Theta} \sum_{i=1}^{m} \left(\rho_i - f(x_i;\Theta)\right)^2 \quad — (1)$$

If $f(x_i;\Theta)$ is linear in Θ, (1) reduces to linear least squares.

Gauss-Newton method for NLLS

given starting guess for $x^{(0)}$

repeat

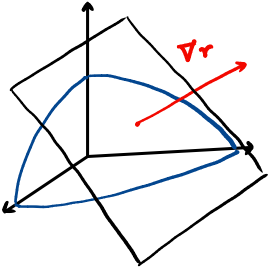
1. linearize r near current guess for $\bar{x} = x^{(k)}$
2. solve a linear LS problem for next step

until converged

Gradient

The gradient of a differentiable function $r : \mathbb{R}^n \to \mathbb{R}$ at x is:

$$\nabla r(x) = \left(\frac{\partial r}{\partial x_1}(x), \frac{\partial r}{\partial x_2}(x), \dots, \frac{\partial r}{\partial x_n}(x) \right) \in \mathbb{R}^n$$



Linearization of $r(x)$ at $\bar{x}$ is:

$$\bar{r}(\bar{x}) = r(\bar{x}) + \frac{\partial r}{\partial x_1}(\bar{x})\,(x_1 - \bar{x}_1) + \dots + \frac{\partial r}{\partial x_n}(\bar{x})\,(x_n - \bar{x}_n)$$

$$= r(\bar{x}) + \nabla r(\bar{x})^T (x - \bar{x})$$

Gradient of residual.

$$r(x) = \begin{bmatrix} r_1(x) \\ \vdots \\ r_n(x) \end{bmatrix}$$

$$\min_x \; f(x) = \sum_{i=1}^{m} r_i(x)^2$$

Necessary condition for optimality of x:

$$\frac{\partial f}{\partial x_j}(x) = \sum_{i=1}^{m} 2 r_i(x) \frac{\partial r_i}{\partial x_j}(x) = 2 \nabla r_i(x)^T r(x)$$

So, $$\nabla f(x) = \begin{bmatrix} \frac{\partial r}{\partial x_1}(x) \\ \vdots \\ \frac{\partial r}{\partial x_n}(x) \end{bmatrix} = 2 \begin{bmatrix} \nabla r_1(x)^T \\ \vdots \\ \nabla r_n(x)^T \end{bmatrix} r(x) = 0$$

Derivative matrix.

necessary condition.

If $r(x) = Ax - b$.

$$\nabla f(x) = 0 \iff A^T(Ax - b) = 0$$

normal equation.

Nonlinear least squares

- The NLLS (nonlinear least-squares) problem:

$$\underset{x \in \mathbf{R}^n}{\text{minimize}} \quad \frac{1}{2}\|r(x)\|_2^2, \quad r : \mathbf{R}^n \to \mathbf{R}^m \quad \text{(typically, } m > n\text{)}.$$

- "Residual" vector

$$r(x) = \begin{bmatrix} r_1(x) \\ r_2(x) \\ \vdots \\ r_m(x) \end{bmatrix}, \quad r_i : \mathbf{R}^n \to \mathbf{R}$$

- Reduces to least-squares when r is **affine:**

$$r(x) = Ax - b$$

Linearization of residual

$\min_x \frac{1}{2} \|r(x)\|_2^2 \approx$

Linearization around a point $\bar{x} \in \mathbf{R}^n$:

$\min_x \frac{1}{2} \|A(\bar{x})x - \bar{b}\|_2^2$

$$r(x) = \begin{bmatrix} r_1(x) \\ \vdots \\ r_m(x) \end{bmatrix} \approx \begin{bmatrix} r_1(\bar{x}) + \nabla r_1(\bar{x})^T(x - \bar{x}) \\ \vdots \\ r_m(\bar{x}) + \nabla r_m(\bar{x})^T(x - \bar{x}) \end{bmatrix} = A(\bar{x})x - b(\bar{x})$$

where

← Derivative matrix

$Ax - r$

$$A(\bar{x}) = \begin{bmatrix} \nabla r_1(x)^T \\ \vdots \\ \nabla r_m(x)^T \end{bmatrix} \in \mathbf{R}^{m \times n}, \qquad b(\bar{x}) = \underline{A(\bar{x})\bar{x} - r(\bar{x})} \in \mathbf{R}^m$$

$x^{(0)}$ – initialization

and $A(\bar{x})$ is the **Jacobian** of mapping r at $\bar{x}$.

Linearized least-squares problem used to determine $x^{(k+1)}$

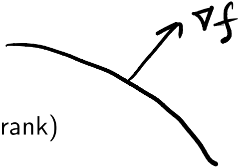
$$\leftarrow x^{(k+1)} = \underset{x \in \mathbf{R}^n}{\operatorname{argmin}} \|A(x^{(k)})x - b(x^{(k)})\|_2^2$$

satisfies $\rightarrow A(x^k)^T A(x^k)\, x = A(x^k)^T b(x^k)$

Dampening

Expand the linear least squares

$$\bar{A}_k = \bar{A} = A(x^{(k)}) \quad \bar{b}_k = b(x^{(k)}), \quad \bar{r}_k = r(x^{(k)})$$



$$\begin{aligned} x^{(k+1)} &= \underset{x \in \mathbf{R}^n}{\operatorname{argmin}} \|\bar{A}x - \bar{b}\|_2^2 \\ &= (\bar{A}^T\bar{A})^{-1}\bar{A}^T\bar{b} \qquad \text{(if } \bar{A} \text{ has full rank)} \\ &= (\bar{A}^T\bar{A})^{-1}\bar{A}^T(\bar{A}x^{(k)} - \bar{r}) \\ &= x^{(k)} - \underbrace{(\bar{A}^T\bar{A})^{-1}\bar{A}^T\bar{r}}_{\text{step}} \end{aligned}$$

d is a descent direction if $d^T\nabla f < 0$

Dampened Gauss-Newton

$$x^{(k+1)} = x^{(k)} - \alpha z^{(k)}, \qquad z^{(k)} = \underset{x \in \mathbf{R}^n}{\operatorname{argmin}} \|A(x^{(k)})x - r(x^{(k)})\|^2$$

for $0 < \alpha \leq 1$.

$$\min_{\alpha > 0} f(x^k + \alpha z^k)$$

Gauss-Newton method for NLLS

$$\min_x \quad \frac{1}{2} \|r(x)\|_2^2$$

given starting guess for $x^{(0)}$

repeat

1. linearize r near current guess for $\bar{x} = x^{(k)}$

$$r(x) \approx r(\bar{x}) - A(\bar{x})(x - \bar{x})$$

2. solve a linear LS problem for next step

$$z^{(k)} = \operatorname*{argmin}_{x \in \mathbf{R}^n} \|A(\bar{x})x - r(\bar{x})\|_2^2$$

3. take damped step

$$x^{(k+1)} = x^{(k)} - \alpha^{(k)} z^{(k)}, \quad 0 < \alpha^{(k)} \leq 1$$

until converged