Least Squares Problems

- Data fitting (or parameter estimation) is an important technique used for modeling in many areas of disciplines.
  - ◇ Assuming a physical phenomenon is modeled by a relationship

$$y = f(z; x_1, \ldots, x_n). \tag{0.1}$$

    - ▷ f is a prescribed function determined up to values of $x_1, \ldots, x_n$.
    - ▷ z is the control variable.
    - ▷ y is the expected response to z.
  - ◇ After m experiments ($m \geq n$), we have collected m observed quantities $(z_i, y_i), i = 1, \ldots, m$.
    - ▷ Due to measurement errors (called noise), (z_i, y_i) may not satisfy (0.1) exactly.
    - ▷ Seek to adjust the parameters $x_1, \ldots, x_n$ so that the expression

$$g(x_1, \ldots, x_n) := \sum_{i=1}^{m} \|y_i - f(z_i; x_1, \ldots, x_n)\|^2 \tag{0.2}$$

      is minimized.
  - ◇ When the norm used in (0.2) is either the 2−norm or the Frobenius norm, we say we have a *least squares problem* .
- An example: Polynomial least squares fitting.
  - ◇ Suppose the function f in (0.1) is an $(n-1)$-th degree polynomial

$$f(z; x_1, \ldots, x_n) = x_1 z^{n-1} + \ldots + x_{n-1} z + x_n. \tag{0.3}$$

◇ The data fitting problem is to solve the system

$$\begin{bmatrix} z_1^{n-1} & z_1^{n-2} & \dots & z_1 & 1 \\ z_2^{n-1} & & & & \\ \vdots & & & & \\ & & & & \\ z_m^{n-1} & z_m^{n-1} & \dots & z_m & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ \\ x_n \end{bmatrix} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ \\ y_m \end{bmatrix} \tag{0.4}$$

for the coefficients $(x_1, \dots, x_n)$.

◇ The system (0.4) is overdetermined, so generally there is no solution.

◇ Instead, we consider the linear least squares problem

$$\min_{x \in R^n} \|Ax - b\|_2^2 \tag{0.5}$$

where $A \in R^{m \times n}$, $b \in R^m$ are known quantities.

- A solution through the QR decomposition.

◇ There exist an orthogonal matrix $Q \in R^{m \times m}$ and $R = \begin{bmatrix} R_1 \\ \\ 0 \end{bmatrix} \in R^{m \times n}$ with R_1 upper triangular such that

$$Q^T A = R.$$

◇ Recall that orthogonal transformations leave the norm $\|x\|_2$ of a vector x invariant ($\|Vx\|_2 = \sqrt{x^T V^T V x} = \|x\|_2$).

◇ Denoting $Q^T b := [h_1^T, h_2^T]^T$, then

$$\begin{aligned} \|Ax - b\|_2^2 &= \|Q^T(Ax - b)\|_2^2 \\ &= \| \begin{bmatrix} R_1 \\ 0 \end{bmatrix} x - \begin{bmatrix} h_1 \\ h_2 \end{bmatrix} \|_2^2 = \|R_1 x - h_1\|_2^2 + \|h_2\|_2^2. \end{aligned}$$

◇ $\|Ax - b\|_2$ is minimized if x is chosen so that

$$R_1 x = h_1. \tag{0.6}$$