Nonlinear Least Squares

$$\underset{x \in \mathbb{R}^n}{\text{minimize}} \; \|g(Ax) - b\|_2^2$$

example $\quad g(x) = \frac{1}{2}\|\sin(Ax) - b\|_2^2$

denominator layout

$$\partial_x y = \begin{pmatrix} \partial_1 y_1 & \cdots & \partial_1 y_m \\ \vdots & \ddots & \vdots \\ \partial_n y_1 & \cdots & \partial_n y_m \end{pmatrix}; \quad x \in \mathbb{R}^n; \quad y \in \mathbb{R}^m$$

nominator layout

$$\partial_{x^T} y = \begin{pmatrix} \partial_1 y_1 & \cdots & \partial_n y_1 \\ \vdots & \ddots & \vdots \\ \partial_1 y_m & \cdots & \partial_n y_m \end{pmatrix}; \quad x \in \mathbb{R}^n; \; y \in \mathbb{R}^m$$

<u>gradient:</u> $\quad \partial_x g(x) = \partial_x \|\sin(Ax) - b\|_2^2, \; x \in \mathbb{R}^n, b \in \mathbb{R}^n, \; A \in \mathbb{R}^{n,n}$

$$\partial_x \|y(x)\|_2^2 = 2\left(\partial_x y(x)\right) y(x) \qquad \text{(denominator layout)}$$

$$\partial_y \sin(y) = \begin{pmatrix} \partial_1 \sin(y_1) & \cdots & \partial_1 \sin(y_n) \\ \vdots & \ddots & \vdots \\ \partial_n \sin(y_1) & \cdots & \partial_n \sin(y_n) \end{pmatrix} = \;?$$

$$\partial_x Ax = \begin{pmatrix} \partial_1 \sum_{i=1}^n a_{1i} x_i & \cdots & \partial_1 \sum_{i=1}^n a_{ni} x_i \\ \vdots & \ddots & \\ \partial_n \sum_{i=1}^n a_{1i} x_i & \cdots & \partial_n \sum_{i=1}^n a_{ni} x_i \end{pmatrix} = \;?$$

$$\Rightarrow \partial_x g(x) = \;?$$

<u>Hessian:</u> $\quad \partial_{x^T} \partial_x g(x) = \partial_{x^T}$

NONLINEAR LEAST SQUARES (MATRIX)

$$\underset{x \in \mathbb{R}^n}{\text{minimize}} \ \frac{1}{2} \| g(AX) - B \|_F^2 = \frac{1}{2}\left(\sqrt{tr(QQ^T)}\right)^2 = \frac{1}{2} tr(QQ^T)$$

$$Q := g(AX) - B$$

$$y : \mathbb{R}^{m,n} \to \mathbb{R}, \qquad \partial_X y(X) = \begin{pmatrix} \partial_{11} y(X) & \partial_{21} y(X) & \cdots & \partial_{m1} y(X) \\ \partial_{12} y(X) & \partial_{22} y(X) & \cdots & \partial_{m2} y(X) \\ \vdots & \vdots & \ddots & \vdots \\ \partial_{1n} y(X) & \partial_{2n} y(X) & \cdots & \partial_{mn} y(X) \end{pmatrix}$$

numerator layout

$$tr(AB) = tr\left(\begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix} (b_1 \ b_2 \cdots b_n) \right) = tr \begin{pmatrix} a_1^T b_1 & a_1^T b_2 & \cdots & a_1^T b_n \\ a_2^T b_1 & a_2^T b_2 & \cdots & a_2^T b_n \\ \vdots & \vdots & \ddots & \vdots \\ a_n^T b_1 & a_n^T b_2 & \cdots & a_n^T b_n \end{pmatrix}$$

$$= \ ?$$

$$=$$

$$tr\left((g(AX) - B)(g(AX) - B)^T \right)$$

$$AX = (Ax_1 \ Ax_2 \cdots Ax_n) = \begin{pmatrix} \sum_{i=1}^n a_{1i} x_{i1} & \sum_{i=1}^n a_{1i} x_{i2} & \cdots & \sum_{i=1}^n a_{1i} x_{in} \\ \sum_{i=1}^n a_{2i} x_{i1} & \sum_{i=1}^n a_{2i} x_{i2} & \cdots & \sum_{i=1}^n a_{2i} x_{in} \\ \vdots & \vdots & \ddots & \vdots \\ \sum_{i=1}^n a_{ni} x_{i1} & \sum_{i=1}^n a_{ni} x_{i2} & \cdots & \sum_{i=1}^n a_{ni} x_{in} \end{pmatrix}$$

$$g(AX) = \begin{pmatrix} g(\sum_{i=1}^n a_{1i} x_{i1}) & g(\sum_{i=1}^n a_{1i} x_{i2}) & \cdots & g(\sum_{i=1}^n a_{1i} x_{in}) \\ g(\sum_{i=1}^n a_{2i} x_{i1}) & g(\sum_{i=1}^n a_{2i} x_{i2}) & \cdots & g(\sum_{i=1}^n a_{2i} x_{i2}) \\ \vdots & \vdots & \ddots & \vdots \\ g(\sum_{i=1}^n a_{ni} x_{i1}) & g(\sum_{i=1}^n a_{ni} x_{i2}) & \cdots & g(\sum_{i=1}^n a_{ni} x_{in}) \end{pmatrix}$$

$$g(AX)-B=\begin{pmatrix} g\left(\sum_{i=1}^{n} a_{1i}x_{i1}\right)-b_{11} & g\left(\sum_{i=1}^{n} a_{1i}x_{i2}\right)-b_{12} & \cdots & g\left(\sum_{i=1}^{n} a_{1i}x_{in}\right)-b_{1n} \\ g\left(\sum_{i=1}^{n} a_{2i}x_{i1}\right)-b_{21} & g\left(\sum_{i=1}^{n} a_{2i}x_{i2}\right)-b_{22} & \cdots & g\left(\sum_{i=1}^{n} a_{2i}x_{i2}\right)-b_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ g\left(\sum_{i=1}^{n} a_{ni}x_{i1}\right)-b_{n1} & g\left(\sum_{i=1}^{n} a_{ni}x_{i2}\right)-b_{n2} & \cdots & g\left(\sum_{i=1}^{n} a_{ni}x_{in}\right)-b_{nn} \end{pmatrix}$$

$$= \begin{pmatrix} q_1 \\ q_2 \\ \vdots \\ q_n \end{pmatrix}$$

$$tr\left(\left(g(AX)-B\right)\left(g(AX)-B\right)^T\right) = \left(q_1^T q_1 + q_2^T q_2 + \cdots + q_n^T q_n\right)$$

$$= \sum_{g=1}^{n}\left(g\left(\sum_{i=1}^{n} a_{1i}x_{ig}\right)-b_{1g}\right)^2 + \sum_{g=1}^{n}\left(g\left(\sum_{i=1}^{n} a_{2i}x_{ig}\right)-b_{2g}\right)^2 + \cdots + \sum_{g=1}^{n}\left(g\left(\sum_{i=1}^{n} a_{ni}x_{ig}\right)-b_{ng}\right)^2$$

---

a $\partial_{11}\left(q_1^T q_1 + q_2^T q_2 + \cdots + q_n^T q_n\right) = ?$

$\partial_{12}\left(q_1^T q_1 + q_2^T q_2 + \cdots + q_n^T q_n\right) = ?$

$\partial_{21}\left(q_1^T q_1 + q_2^T q_2 + \cdots + q_n^T q_n\right) = ?$

$\vdots$

summary

$\partial_{k\ell}\left(q_1^T q_1 + q_2^T q_2 + \cdots + q_n^T q_n\right) = ?$

$$\partial_X y(X) = \begin{pmatrix} \partial_{11} y(X) & \partial_{21} y(X) & \cdots & \partial_{n1} y(X) \\ \partial_{12} y(X) & \partial_{22} y(X) & \cdots & \partial_{n2} y(X) \\ \vdots & \vdots & \ddots & \vdots \\ \partial_{1n} y(X) & \partial_{2n} y(X) & \cdots & \partial_{nn} y(X) \end{pmatrix} = Z$$

$$= \begin{pmatrix} z_{11} & z_{12} & \cdots & z_{1n} \\ z_{21} & z_{22} & \cdots & z_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ z_{n1} & z_{n2} & \cdots & z_{nn} \end{pmatrix}$$

$$z_{11} = ?$$

$$z_{21} = ?$$

$$z_{12} = ?$$

<u>Overview</u>

$$z_{\ell k} = \partial_{k\ell}\, y(X) = ?$$