steepest α 相

$$\frac{d}{dx}|Ax_0 - b|^2 = g(0) = 2A^TAx - 2A^Tb = 2A^T(Ax-b)$$

$$\frac{d}{d\alpha} f(1) = g(x_1) \cdot g(0)$$

$$= [g(x_1)]^T \times g(0)$$

$$= (2A^TAx_1 - 2A^Tb)^T \cdot g(0) \qquad \longrightarrow x_0 + \alpha \cdot g(0)$$

$$= (2A^TAx_0 + 2A^T\alpha \cdot g(0) - 2A^Tb)^T g(0)$$

$$= \boxed{2A^T(Ax_0 - b) + 2\alpha A^TA g(0)}^T g(0)$$

$$= [\; g(0) + 2\alpha A^TA g(0)\;]^T g(0)$$

$$0 = g(0)^T g(0) + 2\alpha g(0)^T A^TA g(0)$$

$$\alpha = \frac{g(0)^T g(0)}{g(0)^T A^TA g(0)} \qquad \text{learning rate update}$$

learning rate   gradient

$$x_1 = x_0 - \alpha \cdot g(0)$$

$$g(0) = \frac{d}{dx}|Ax-b|^2$$
$$= 2A^TAx - 2A^Tb$$