$$\theta' = \underbrace{\theta^0}_{\text{起点}} + \underbrace{\alpha}_{\text{步长}} \underbrace{\nabla J(\theta)}_{\text{损失函数}}$$

多变量：

(方向向量)

$J(\theta_1, \theta_2) = \theta_1^2 + \theta_2^2$

$J'_{\theta_1} = 2\theta_1$

$J'_{\theta_2} = 2\theta_2$

$\nabla J(\theta_1, \theta_2) = \langle 2\theta_1, 2\theta_2 \rangle$

令 $\theta_0 = (1, 3)$, $\alpha = 0.1$

$\theta_1 = \theta_0 - 0.1 \times (2, 6) = (0.8, 2.4)$

$\theta_2 = (0.8, 2.4) - 0.1 \times (1.6, 4.8)$

$= (0.64, 1.92)$

$\vdots$

单变量

(斜率)

$J(\theta) = \theta^2$

$J'(\theta) = 2\theta$

令 $\theta_0 = 1$, $\alpha = 0.2$

$\theta_1 = \theta_0 - \alpha \boxed{\nabla J(\theta_0)}$

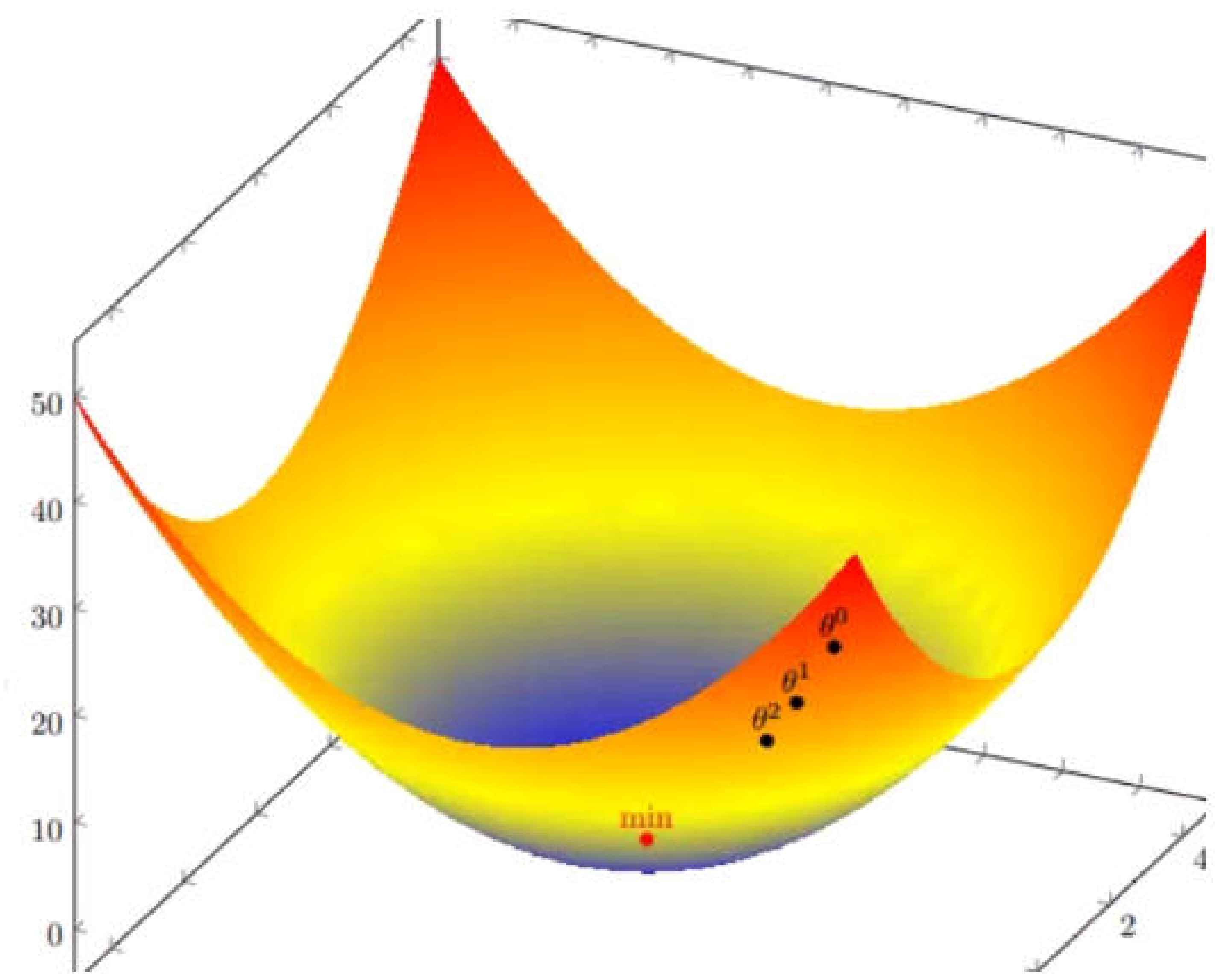
$= 1 - 0.2 \times 2 = 0.6$

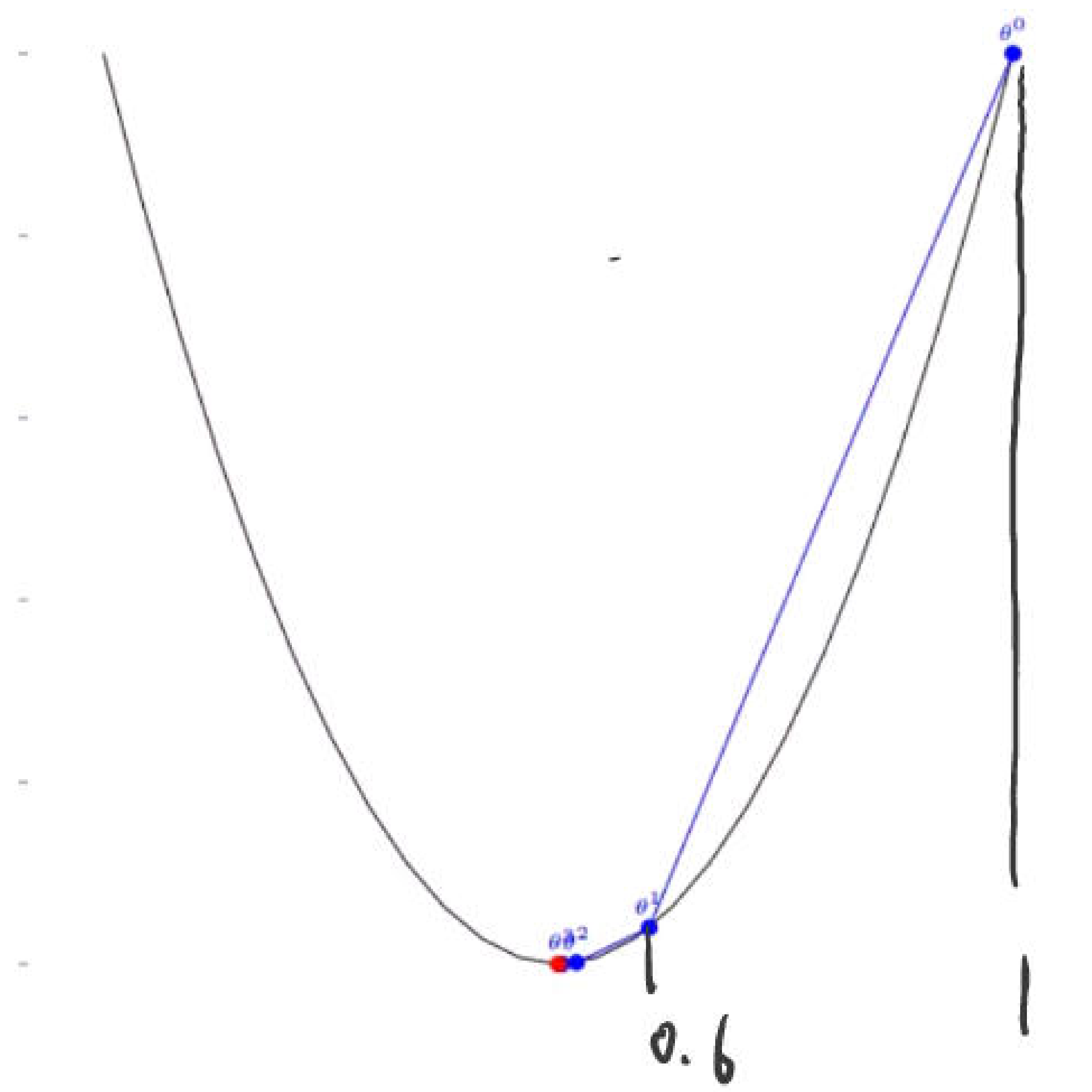
$\theta_2 = \theta_1 - \alpha \nabla J(\theta_1)$

$= 0.6 - 0.2 \times 1.2 = 0.36$

$\vdots$

迭代





用梯度下降法实现线性回归

损失函数：

预测函数

$$J(\theta)=\frac{1}{2m}\sum_{i=1}^{m}(\boxed{h_\theta(x_i)}-y_i)^2$$

$$h_\theta(x_i)=\theta_0+\theta_1 x_1+\theta_2 x_2+\cdots+\theta_i x_i$$

$$=\sum_{i=0}^{i}\theta_i x_i$$

可以是 x_i 这个特征元素，$x_i^{(i)}$

$$h_\theta(x)=\theta_0+\theta_1 x$$

$$f(x)=b+ax$$

已知 x、$h_\theta(x)$，求 θ

$$h_\theta(x_i)=\theta_0+\theta_1 x_i$$

(x_i, y_i) 是一个真实点

$h_\theta(x_i)$ 是预测点

$$\nabla J(\theta)=\left\langle\frac{\partial J}{\partial\theta_0},\frac{\partial J}{\partial\theta_1}\right\rangle$$

$$\frac{1}{m}\sum_{i=1}^{m}(h_\theta(x_i)-y_i)$$

$$\frac{1}{m}\sum_{i=1}^{m}(h_\theta(x_i)-y_i)x_i$$

$$J(\theta)=\frac{1}{2m}\sum_{i=1}^{m}\left(h_\theta(x_i)-y_i\right)^2$$

令: $h_\theta(x)=\theta_0+\theta_1 x_1^{(i)}=X\theta$

两个变量

$\theta_0 x_0+\theta_1 x_1^{(i)}$

X增加一维 x_0，值固定为1，乘到θ_0上。

$$\vec{x}=(\overset{1}{\overset{\|}{x_0}},x_1^{(i)})$$

$$\vec{\theta}=\begin{pmatrix}\theta_0\\ \theta_1\end{pmatrix}$$

$$J(\theta)=\frac{1}{2m}(\vec{x}\vec{\theta}-\hat{y})^T(X\theta-\hat{y})$$

$$\nabla J(\theta)=\frac{1}{m}X^T(\vec{x}\vec{\theta}-\hat{y})$$

$$Y=\begin{bmatrix}3\\4\\5\\5\\2\\4\\\vdots\\21\end{bmatrix} \text{20个数}$$

$$X_0=\begin{bmatrix}1\\1\\\vdots\\1\end{bmatrix}\Big\} m个$$

$$X_1=\begin{bmatrix}1\\2\\3\\4\\\vdots\\m\end{bmatrix} m个$$

$$X_2=\begin{bmatrix}1,1\\1,2\\1,3\\\vdots\\1,4\\\vdots\\1,m\end{bmatrix}$$

$$\vec{A} = \vec{X}\vec{\theta} - \hat{y} =$$

$$\underset{m\times 2}{\begin{bmatrix} 1,1 \\ 1,2 \\ 1,3 \\ \vdots \\ 1,m \end{bmatrix}} \times \underset{2\times 1}{\begin{bmatrix} \theta_0 \\ \theta_1 \end{bmatrix}} = \underset{m\times 1}{\begin{bmatrix} \theta_0+\theta_1 \\ \theta_0+2\theta_1 \\ \theta_0+3\theta_1 \\ \vdots \\ \theta_0+m\theta_1 \end{bmatrix}} - \begin{bmatrix} \\ \\ \\ \end{bmatrix}$$

$$\vec{A}\cdot\vec{A}^T = \text{一个数}$$

$$y = \begin{bmatrix} 3 \\ 4 \\ 5 \\ 5 \\ 2 \\ 4 \\ \vdots \\ 21 \end{bmatrix} \quad \text{20个数}$$

4

$$X_0 = \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix} \Big\} m\text{个}$$

$$X_1 = \begin{bmatrix} 1 \\ 2 \\ 3 \\ 4 \\ \vdots \\ m \end{bmatrix} \quad m\text{个}$$

$$\vec{X} = \begin{bmatrix} 1,1 \\ 1,2 \\ 1,3 \\ \vdots \\ \vdots,4 \\ \vdots \\ 1,m \end{bmatrix}$$

$$m\times 2 \quad 2\times 1 = m\times 1$$