线性回归的一般形式：

$$f(\boldsymbol{x}) = \boldsymbol{w}^T \boldsymbol{x} + b$$

损失函数的一般形式：

$$L(\theta) = \frac{1}{2m} \sum_{i=1}^{m} [f(\boldsymbol{x_i}) - y_i]^2$$

使用正规方程求解：

$$\boldsymbol{X} = \begin{bmatrix} x_1^{(1)} & \cdots & 1 \\ \vdots & \ddots & \vdots \\ x_1^{(m)} & \cdots & 1 \end{bmatrix} m\text{个样本，}n\text{个特征，并在最后添加一列}1$$

$$\boldsymbol{Y} = \begin{bmatrix} y^{(1)} \\ \vdots \\ y^{(m)} \end{bmatrix}$$

$$\boldsymbol{\theta} = \begin{bmatrix} \theta_1 \\ \vdots \\ \theta_{n+1} \end{bmatrix}$$

L2 正则化项：λ=0 就是不带正则化项

$$L(\theta) = \sum_{i=1}^{m} [f(\boldsymbol{x_i}) - y_i]^2 + \lambda \sum_{j=1}^{n} \theta_j^2 = \theta^T X^T X \theta - Y^T X \theta + Y^T Y - \theta^T X^T Y + \lambda (E\theta)^T E\theta$$

$$\nabla_\theta L = 2X^T X\theta - 2X^T Y + 2\lambda E\theta = \boldsymbol{0} \Rightarrow \boldsymbol{\theta} = (X^T X + \lambda E)^{-1} X^T Y$$

注意，这里的 E 不是单位矩阵，而是：

$$E = \begin{bmatrix} 1 & \cdots & 0 \\ \vdots & 1 & \vdots \\ 0 & \cdots & 0 \end{bmatrix} (n+1, n+1)$$ 最后一个元素为 0（对应 X 最后一列为常数项 1）

使用梯度下降法来求解：

$$X=\begin{bmatrix} x_0^{(1)} & \cdots & x_n^{(1)} \\ \vdots & \cdots & \vdots \\ x_0^{(m)} & \cdots & x_n^{(m)} \end{bmatrix}(m,n+1)\text{首列为常数项 1}$$

$$\boldsymbol{Y}=\begin{bmatrix} y^{(1)} \\ \vdots \\ y^{(m)} \end{bmatrix}$$

$$\boldsymbol{\theta}=\begin{bmatrix} \theta_0 \\ \vdots \\ \theta_n \end{bmatrix}\theta_0\text{为截距项}$$

$$f(x)=x\theta$$

$$L(\theta)=\frac{1}{2m}\sum_{i=1}^{m}[f(\boldsymbol{x_i})-y_i]^2+\frac{\lambda}{2}\sum_{j=1}^{n}\theta_j^2=\frac{1}{2m}(\theta^TX^TX\theta-Y^TX\theta+Y^TY-\theta^TX^TY)+\frac{\lambda}{2}(E\theta)^TE\theta$$

$$\nabla_\theta L=\frac{1}{m}X^T(X\theta-Y)+\lambda E\theta$$