

1、对于任意 n 个点二维 (x, y) ，我们都可以用插值法进行插值，得到一个 $n - 1$ 次的曲线

对于 n 个点 (x_i, y_i) 来说，可以用如下公式

$$f(x) = \sum_{i=1}^n \frac{\prod_{j \neq i} (x - x_j)}{\prod_{j \neq i} (x_i - x_j)} y_i$$

多维数据可以简单进行扩展。

如果需要进行额外修正，可以增加高次额外项 $g(x) = \prod_{i=1}^n (x - x_i) h(x)$ 满足要求。

2、不妨按照书中，向量 \vec{x} 都是列向量。

标量 y 对向量 \vec{x} 求导

$$\frac{\partial y}{\partial \vec{x}} = \left(\frac{\partial y}{\partial x_1}, \dots, \frac{\partial y}{\partial x_n} \right)$$

于是乎 $\frac{\partial \vec{x}^T \vec{y}}{\partial \vec{x}} = \vec{y}$, $(\vec{x}^T \vec{y} = \vec{y}^T \vec{x})$

线性回归:(此时为标量 x)

$$(\omega^*, b^*) = \arg \min_{(\omega, b)} \sum_{i=1}^m (f(x_i) - y_i)^2 \quad (1)$$

$$= \arg \min_{(\omega, b)} \sum_{i=1}^m (y_i - \omega x_i - b)^2 \quad (2)$$

$$E(\omega, b) = \sum_{i=1}^m (y_i - \omega x_i - b)^2$$

对 ω, b 分别求偏导数：(这里不是很清楚如何判断多参数的函数是否为凸函数，但是这里单独看每个参数，都是一个二次函数，因此偏导为 0 的点就是最优)

$$\frac{\partial E}{\partial \omega} = 2 \sum_{i=1}^m (\omega x_i^2 + x_i(b - y_i))$$

$$\frac{\partial E}{\partial b} = 2 \sum_{i=1}^m (b + (\omega x_i - y_i))$$

所以我们有：

$$b = \frac{\sum_{i=1}^m (y_i - \omega x_i)}{m} = \bar{y} - \omega \bar{x}$$

$$\begin{aligned}
0 &= \sum_{i=1}^m (\omega x_i^2 + x_i(b - y_i)) \\
&= \sum_{i=1}^m (\omega x_i^2 + x_i(\bar{y} - \omega \bar{x} - y_i)) \\
&= \sum_{i=1}^m \omega x_i^2 + \bar{y} * \sum_{i=1}^m x_i + \omega \bar{x} \sum_{i=1}^m x_i - \sum_{i=1}^m x_i y_i \\
&= \sum_{i=1}^m \omega x_i^2 + m\bar{x}\bar{y} + m\omega\bar{x}^2 - \sum_{i=1}^m x_i y_i
\end{aligned}$$

因此有：

$$\omega = \frac{\sum_{i=1}^m x_i y_i - m\bar{x}\bar{y}}{\sum_{i=1}^m x_i^2 - m\bar{x}^2}$$

其中：

$$\begin{aligned}
&\sum_{i=1}^m x_i y_i - m\bar{x}\bar{y} \\
&= \sum_{i=1}^m (x_i - \bar{x})(y_i - \bar{y}) \\
&= \sum_{i=1}^m x_i (y_i - \bar{y}) \\
&= \sum_{i=1}^m (x_i - \bar{x}) y_i
\end{aligned}$$

向量：

这里可以令 $\hat{\omega} = [\omega; b]$, $X_i = [x_i^T \ 1]$ 那么 $X = [X_1; X_2; \dots; X_n]$

变为优化 $\hat{\omega}^* = \arg \min_{\hat{\omega}} \|X\hat{\omega} - y\|$

$$E(\hat{\omega}) = \|X\hat{\omega} - y\|^2 = (X\hat{\omega} - y)^T (X\hat{\omega} - y)$$

求偏导数：

$$\frac{\partial E}{\partial \hat{\omega}} = 2X^T(X\hat{\omega} - y)$$

当 X 为满秩矩阵时，有逆，因此：

$$\hat{\omega} = (X^T X)^{-1} X^T y$$

对数线性回归：令 $y' = \ln y$ 然后求解。