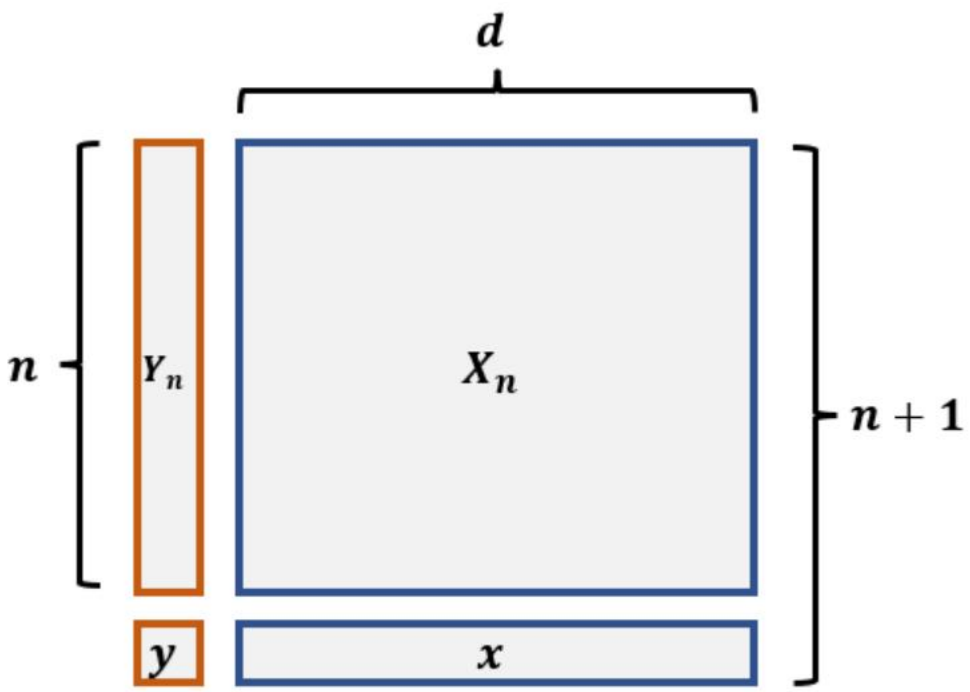
# 作业二

题目：对于数据$X_n \in R^{n\times d}$，标签$Y_n \in R^{n\times 1}$，我们可知其线性参数解为：

$$W_n = (X_n^T X_n)^{-1} X_n^T Y_n$$

其中$X_n^T X_n$为可逆矩阵

问题 1：如图所示，若在原数据基础上增添一个样本（数据为 $x \in R^{1\times d}$，标签为 $y \in R$），得数据$X_{n+1} \in R^{(n+1)\times d}$,标签$Y_{n+1} \in R^{(n+1)\times 1}$，试用$W_n + \alpha x^T(xW_n - y)$的形式来表示$W_{n+1}$（求α）。




**提示**：1.使用分块矩阵求解

## 2. 使用 Sherman-Morrison-Woodbury 公式求解

Sherman-Morrison-Woodbury 公式：

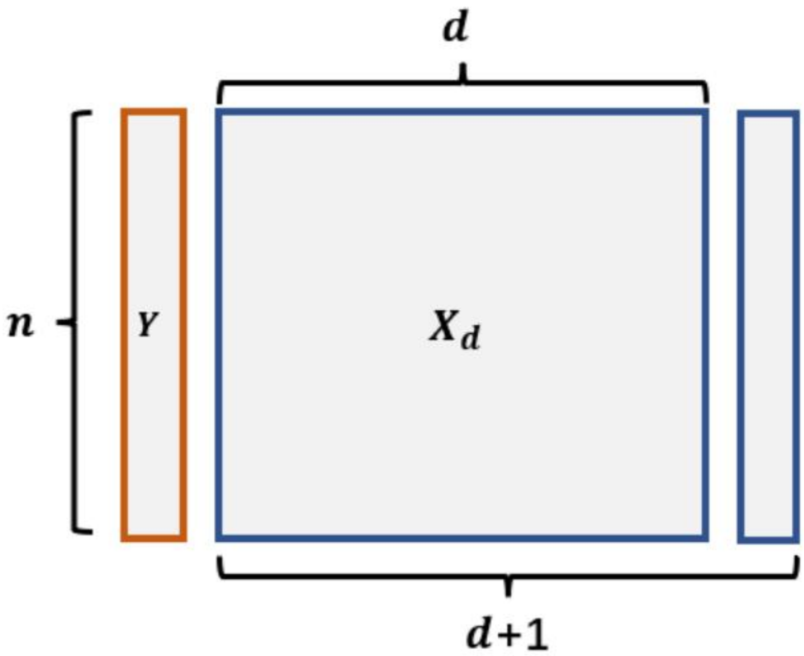
设 A 为 n 阶可逆方阵且$A^{-1}$已知，$B = A + XRY$，其中 X 为 n×r 矩阵，R 为 r 阶可逆方阵，Y 为 r×n 矩阵，则

$$\mathbf{B^{-1} = A^{-1} - A^{-1}X(R^{-1} + YA^{-1}X)^{-1}YA^{-1}}$$

其意义在于当 r≪n 时求 R 与$(R^{-1} + YA^{-1}X)$的逆矩阵远易于直接求 B 的逆矩阵。在本问题中，R 为实数 1，$X \in R^{n\times 1}$, $Y \in R^{1\times n}$。

问题 2：如图所示，若在原数据基础上($X_d \in R^{n\times d}$,标签$Y_d \in$

$R^{n\times 1}$)增添一维特征，得数据$X_{d+1} \in R^{n\times(d+1)}$，标签不变，试用$W_d$加增量的形式表示$W_{d+1}$




**提示**：1.使用分块矩阵求解

$$X_{d+1} = (X_d|x)$$

$$X_{d+1}^T X_{d+1} = \begin{pmatrix} X_d^T \\ x^T \end{pmatrix} (X_d|x) = \begin{pmatrix} X_d^T X_d & X_d^T x \\ x^T X_d & x^T x \end{pmatrix}$$

2.使用四分块矩阵求逆公式

（参考链接 [https://www.zhihu.com/question/47760591）](https://www.zhihu.com/question/47760591)

3. 最终形式与卡尔曼滤波类似

问题 3:若在原数据基础上($X_d \in R^{n\times d}$,标签$Y_d \in R^{n\times 1}$)先增添一个样本（数据为$x_{n+1} \in R^{1\times d}$，标签为 $y \in R$），再增

添一维特征（所增添特征$x_{d+1} \in R^{(n+1)\times 1}$），得数据$X_{n+1,d+1} \in R^{(n+1)\times(d+1)}$，标签$Y_{n+1} \in R^{(n+1)\times 1}$，试用$W_{n,d}$加增量的形式表示$W_{n+1,d+1}$

**提示**：使用在前两个问题中所得到的结论

解答：

问题 1：

由 S-M-W 公式，

$$(X_{n+1}^T X_{n+1})^{-1} = (X_n^T X_n)^{-1} - \frac{(X_n^T X_n)^{-1} x^T x (X_n^T X_n)^{-1}}{1 + x(X_n^T X_n)^{-1} x^T}$$

由于$W_{n+1} = (X_{n+1}^T X_{n+1})^{-1} X_{n+1}^T Y_{n+1}$

其中$X_{n+1}^T Y_{n+1} = \begin{pmatrix} X_n^T & x^T \end{pmatrix} \begin{pmatrix} Y_n \\ y \end{pmatrix} = X_n^T Y_n + x^T y$

所以$W_{n+1} = (X_{n+1}^T X_{n+1})^{-1} \ (X_n^T Y_n + x^T y)$

$=(X_{n+1}^T X_{n+1})^{-1} [(X_n^T X_n)(X_n^T X_n)^{-1} X_n^T Y_n + x^T y]$

$=(X_{n+1}^T X_{n+1})^{-1} [(X_n^T X_n) W_n + x^T y]$

由分块矩阵乘法可知，

$$X_{n+1}^T X_{n+1} = \begin{pmatrix} X_n^T & x^T \end{pmatrix} \begin{pmatrix} X_n \\ x \end{pmatrix} = X_n^T X_n + x^T x$$

故$W_{n+1} = (X_{n+1}^T X_{n+1})^{-1}[(X_{n+1}^T X_{n+1} - x^T x)W_n + x^T y]$

$=W_n - (X_{n+1}^T X_{n+1})^{-1} x^T x W_n + (X_{n+1}^T X_{n+1})^{-1} x^T y$

$=W_n - (X_{n+1}^T X_{n+1})^{-1} x^T \ (xW_n - y)$

对照题目形式，可知$\alpha \ = -(X_{n+1}^T X_{n+1})^{-1}$

$$= \frac{(X_n^T X_n)^{-1} x^T x (X_n^T X_n)^{-1}}{1 + x(X_n^T X_n)^{-1} x^T} - (X_n^T X_n)^{-1}$$

## 问题 2：

问题2：

$$W_{d+1} = (X_{d+1}^T X_{d+1})^{-1} X_{d+1}^T Y$$

$$= \left[ \begin{pmatrix} X_d^T \\ x^T \end{pmatrix} (X_d \;\vdots\; x) \right]^{-1} X_{d+1}^T Y$$

$$= \begin{bmatrix} X_d^T X_d & X_d^T x \\ x^T X_d & x^T x \end{bmatrix}^{-1} \begin{pmatrix} X_d^T Y \\ x^T Y \end{pmatrix}$$

由四分块矩阵求逆公式

$$\begin{bmatrix} X_d^T X_d & X_d^T x \\ x^T X_d & x^T x \end{bmatrix}^{-1} =$$

$$\begin{bmatrix} \left(X_d^T X_d - \frac{1}{x^T x} X_d^T x x^T X_d\right)^{-1} & (X_d^T X_d)^{-1} X_d^T x \left(x^T X_d (X_d^T X_d)^{-1} X_d^T x - x^T x\right)^{-1} \\ \frac{1}{x^T x} x^T X_d \left(\frac{1}{x^T x} X_d^T x x^T X_d - X_d^T X_d\right)^{-1} & \left(x^T x - x^T X_d (X_d^T X_d)^{-1} X_d^T x\right)^{-1} \end{bmatrix},$$

展开 $W_{d+1}$，有 $W_{d+1}$

$$= \begin{bmatrix} \left(X_d^T X_d - \frac{1}{x^T x} X_d^T x x^T X_d\right)^{-1} X_d^T Y + (X_d^T X_d)^{-1} X_d^T x \left(x^T X_d (X_d^T X_d)^{-1} X_d^T x - x^T x\right)^{-1} x^T Y \\ \frac{1}{x^T x} x^T X_d \left(\frac{1}{x^T x} X_d^T x x^T X_d - X_d^T X_d\right)^{-1} X_d^T Y + \left(x^T x - x^T X_d (X_d^T X_d)^{-1} X_d^T x\right)^{-1} x^T Y \end{bmatrix}, \text{记作} \begin{bmatrix} ① \\ ② \end{bmatrix}$$

由 S-M-W 公式及相关变换，可得

$$W_{d+1} = \begin{pmatrix} \left[ I_d + \frac{A^{-1}BC}{D - CA^{-1}B} \right] W_d - \frac{A^{-1}B}{D - CA^{-1}B} x^T Y \\ -\frac{C}{D} \left[ I_d + \frac{A^{-1}BC}{D - CA^{-1}B} \right] W_d + \frac{1}{D - CA^{-1}B} x^T Y \end{pmatrix},$$

其中 $\begin{bmatrix} A & B \\ C & D \end{bmatrix} = \begin{bmatrix} X_d^T X_d & X_d^T x \\ x^T X_d & x^T x \end{bmatrix}$