第六讲 最小二乘法及其应用

贺丽君
信息与通信工程学院
Email: lijunhe@mail.xjtu.edu.cn
2023-02

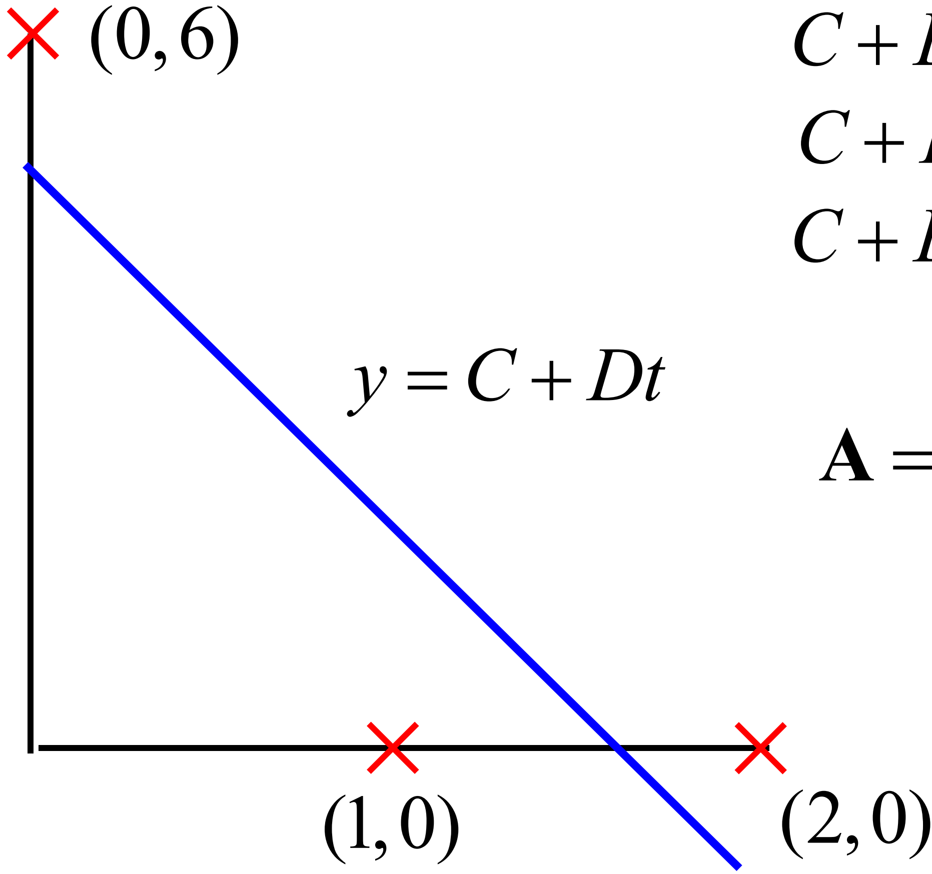
内容提要

- 最小二乘法
- 应用举例1：Netflix电影推荐
- 应用举例2：信道估计

内容提要

- 最小二乘法
- 应用举例1：Netflix电影推荐
- 应用举例2：信道估计

最小二乘法的提出——直线拟合



$$C + D\cdot 0 = 6$$
$$C + D\cdot 1 = 0 \quad \Longrightarrow \quad \mathbf{Ax} = \mathbf{b}$$
$$C + D\cdot 2 = 0$$

$$\mathbf{A} = \begin{bmatrix} 1 & 0 \\ 1 & 1 \\ 1 & 2 \end{bmatrix} \quad \mathbf{x} = \begin{bmatrix} C \\ D \end{bmatrix} \quad \mathbf{b} = \begin{bmatrix} 6 \\ 0 \\ 0 \end{bmatrix}$$

方程组Ax=b无解！

$$\left\| Ax - b \right\|^2$$

$$\downarrow \quad \searrow \quad \searrow$$

$$m\times n \quad n\times 1 \quad m\times 1$$

最小二乘法的提出——直线拟合

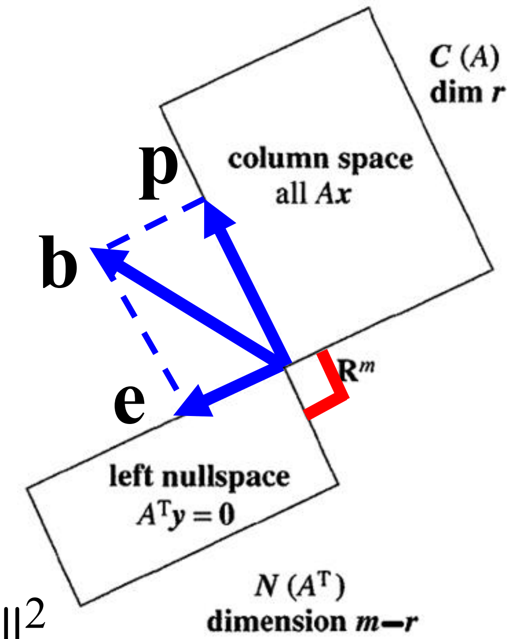
$$\text{minimize } \|\mathbf{Ax}-\mathbf{b}\|^2$$

向量b可以分解为两部分之和：

$$\mathbf{b}=\mathbf{p}+\mathbf{e}$$

列空间分量

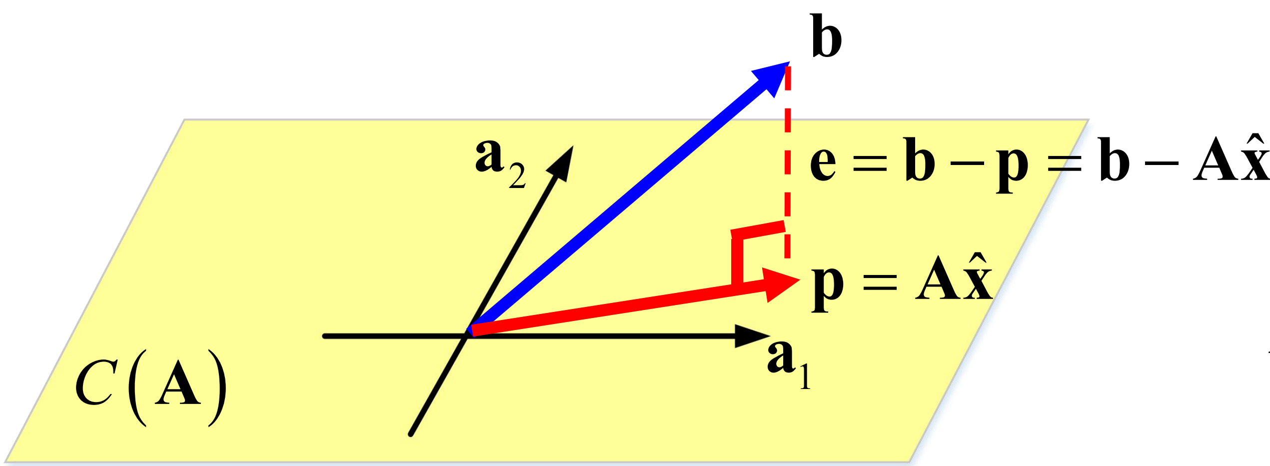
左零空间分量



$$\|\mathbf{Ax}-\mathbf{b}\|^2=\|\mathbf{Ax}-\mathbf{p}-\mathbf{e}\|^2=\underline{\|\mathbf{Ax}-\mathbf{p}\|^2}+\|\mathbf{e}\|^2$$

该项等于零，则目标函数最小

子空间投影



假设各列线性无关

$$\mathbf{A}=\left[\mathbf{a}_1\cdots\mathbf{a}_n\right]$$

$$\underbrace{\begin{bmatrix} — a_1^{\mathrm{T}} — \\ \vdots \\ — a_n^{\mathrm{T}} — \end{bmatrix}}_{\mathbf{A}^T}\begin{bmatrix} b-A\widehat{x} \end{bmatrix}=\begin{bmatrix} \mathbf{0} \end{bmatrix} \Rightarrow$$

$$\hat{\mathbf{x}}=\left(\mathbf{A}^T\mathbf{A}\right)^{-1}\mathbf{A}^T\mathbf{b}$$

$$\mathbf{p}=\mathbf{A}\hat{\mathbf{x}}=\mathbf{A}\left(\mathbf{A}^T\mathbf{A}\right)^{-1}\mathbf{A}^T\mathbf{b}$$

$$\boxed{\mathbf{P}=\mathbf{A}\left(\mathbf{A}^T\mathbf{A}\right)^{-1}\mathbf{A}^T}$$

$$\boxed{\mathbf{P}^2=\mathbf{P},\ \mathbf{P}^T=\mathbf{P}}$$

子空间投影

最小二乘法的提出——直线拟合

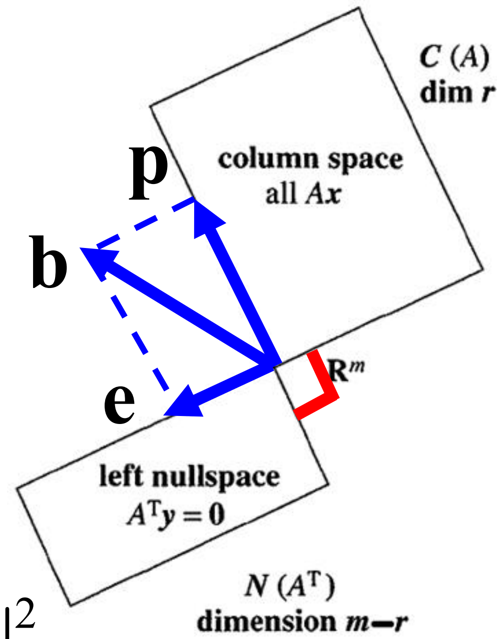
$$\text{minimize } \|\mathbf{Ax}-\mathbf{b}\|^2$$

向量b可以分解为两部分之和：

$$\mathbf{b}=\mathbf{p}+\mathbf{e}$$

列空间分量

左零空间分量

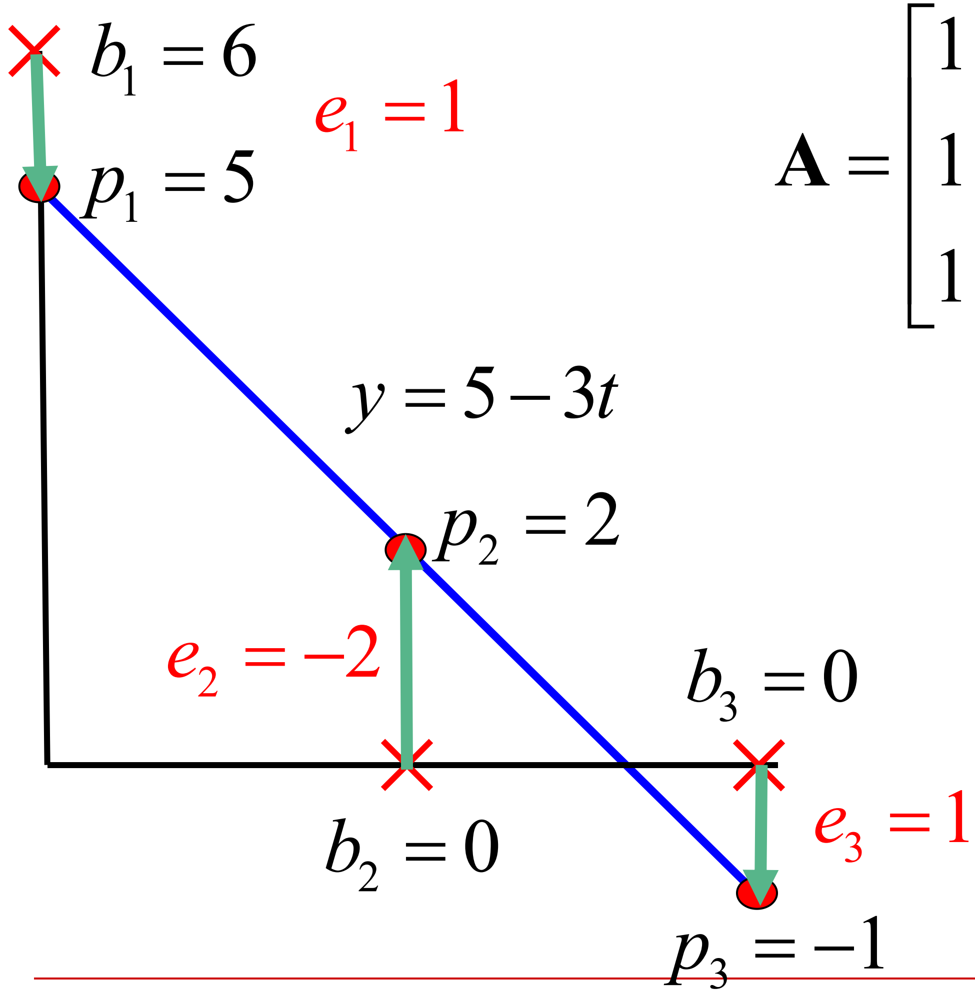


$$\|\mathbf{Ax}-\mathbf{b}\|^2=\|\mathbf{Ax}-\mathbf{p}-\mathbf{e}\|^2=\underline{\|\mathbf{Ax}-\mathbf{p}\|^2}+\|\mathbf{e}\|^2$$

该项等于零，则目标函数最小

$$\Rightarrow \hat{\mathbf{x}}=\left(\mathbf{A}^T\mathbf{A}\right)^{-1}\mathbf{A}^T\mathbf{b}$$ (最小二乘解, Least Squares)

最小二乘法的提出——直线拟合



$$\mathbf{A} = \begin{bmatrix} 1 & 0 \\ 1 & 1 \\ 1 & 2 \end{bmatrix} \quad \mathbf{b} = \begin{bmatrix} 6 \\ 0 \\ 0 \end{bmatrix} \quad \hat{\mathbf{x}} = \begin{bmatrix} 5 \\ -3 \end{bmatrix}$$

$$\|\mathbf{e}\|^2 = e_1^2 + e_2^2 + e_3^2 = 6$$

$$e_1 + e_2 + e_3 = 0$$

为什么?

直线拟合的一般情况

$$\begin{aligned} C + Dt_1 &= b_1 \\ C + Dt_2 &= b_2 \\ &\vdots \\ C + Dt_m &= b_m \end{aligned} \qquad A = \begin{bmatrix} 1 & t_1 \\ 1 & t_2 \\ \vdots & \vdots \\ 1 & t_m \end{bmatrix}$$

$$A^{\mathrm{T}}A = \begin{bmatrix} 1 & \cdots & 1 \\ t_1 & \cdots & t_m \end{bmatrix} \begin{bmatrix} 1 & t_1 \\ \vdots & \vdots \\ 1 & t_m \end{bmatrix} = \begin{bmatrix} m & \sum t_i \\ \sum t_i & \sum t_i^2 \end{bmatrix}$$

$$A^{\mathrm{T}}\boldsymbol{b} = \begin{bmatrix} 1 & \cdots & 1 \\ t_1 & \cdots & t_m \end{bmatrix} \begin{bmatrix} b_1 \\ \vdots \\ b_m \end{bmatrix} = \begin{bmatrix} \sum b_i \\ \sum t_i b_i \end{bmatrix}$$

直线拟合的一般情况

LS本质是参数拟合，不限于直线

标准方程 (Normal Equation)：

$$\begin{bmatrix} m & \sum t_i \\ \sum t_i & \sum t_i^2 \end{bmatrix} \begin{bmatrix} C \\ D \end{bmatrix} = \begin{bmatrix} \sum b_i \\ \sum t_i b_i \end{bmatrix}$$

$$\mathbf{A}^T\mathbf{A} \qquad \hat{\mathbf{x}} \qquad \mathbf{A}^T\mathbf{b}$$

$$E(x) = \|Ax - b\|^2 = (C + Dt_1 - b_1)^2 + \cdots + (C + Dt_m - b_m)^2$$

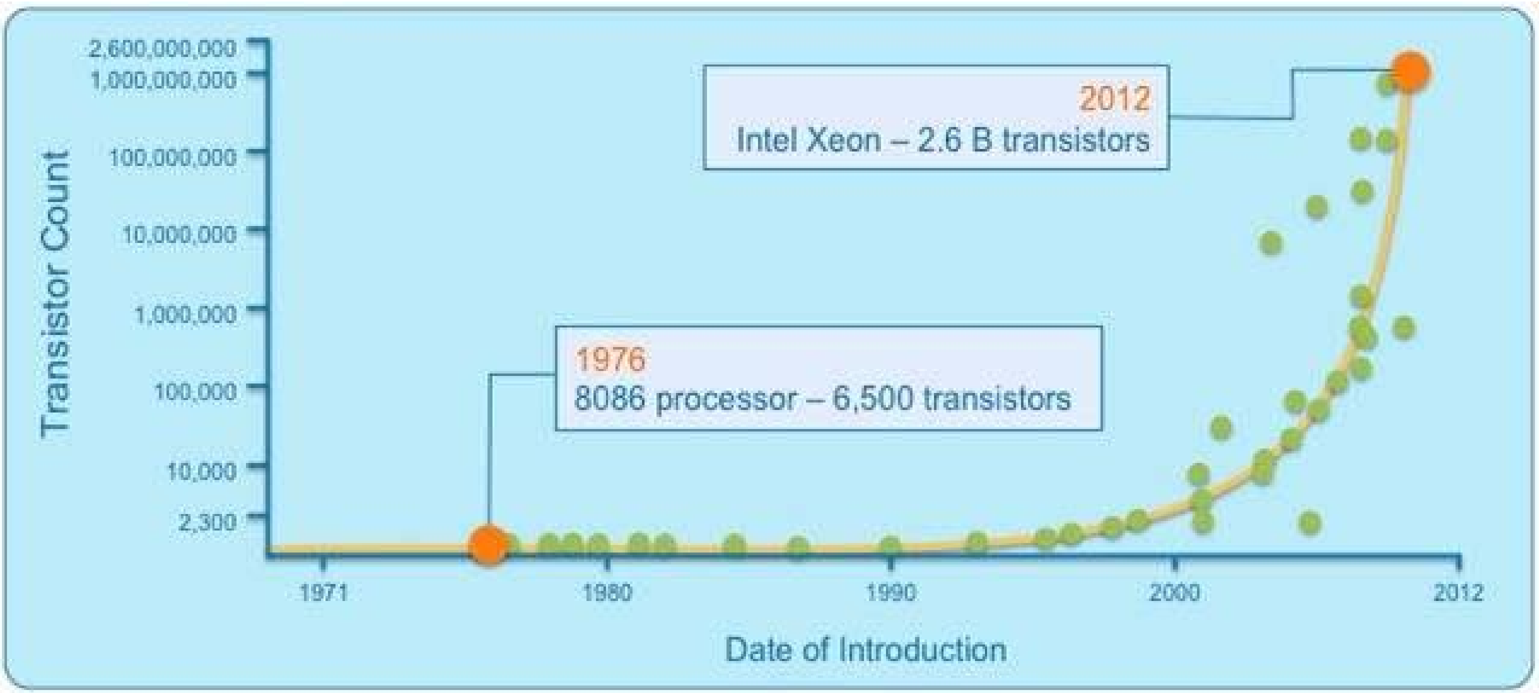
$$\frac{\partial E}{\partial C} = 0 \Rightarrow mC + D\sum t_i = \sum b_i$$

$$\frac{\partial E}{\partial D} = 0 \Rightarrow C\sum t_i + D\sum t_i^2 = \sum t_i b_i$$

与标准方程一致

曲线拟合：摩尔定律

intel价格不变的情况下，处理器包含的晶体管的数量变化规律

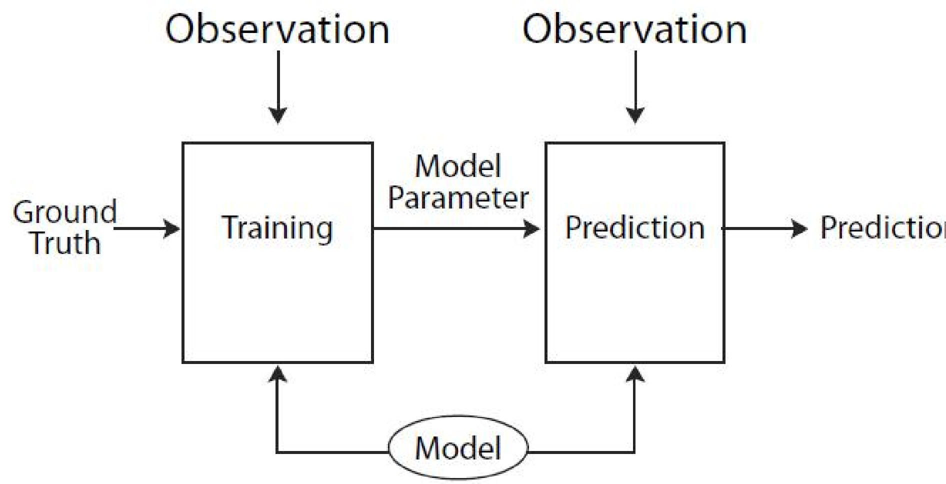


戈登·摩尔的经验之谈：集成电路上可以容纳的晶体管数目在大约每经过**18**个月到**24**个月便会增加一倍。处理器的性能大约每两年翻一倍，同时价格下降为之前的一半。

内容提要

- 最小二乘法
- 应用举例1：Netflix电影推荐
- 应用举例2：信道估计

Netflix电影个性化推荐



电影推荐问题：
预测器设计

Netflix电影个性化推荐

体现用户偏好
和影片质量

$$\hat{r}_{ui} = \bar{r} + \underline{b_u + b_i}$$

平均值=3.83

$$R = \begin{array}{c} \\ 1 \\ 2 \\ 3 \\ 4 \\ 5 \\ 6 \\ 7 \\ 8 \\ 9 \\ 10 \end{array} \begin{array}{c} \begin{array}{ccccc} A & B & C & D & E \end{array} \\ \left(\begin{array}{ccccc} 5 & 4 & 4 & - & \boxed{5} \\ - & 3 & 5 & \boxed{3} & 4 \\ 5 & 2 & - & \boxed{2} & 3 \\ - & \boxed{2} & 3 & 1 & 2 \\ 4 & - & \boxed{5} & 4 & 5 \\ \boxed{5} & 3 & - & 3 & 5 \\ 3 & \boxed{2} & 3 & 2 & - \\ 5 & \boxed{3} & 4 & - & 5 \\ \boxed{4} & 2 & 5 & 4 & - \\ \boxed{5} & - & 5 & 3 & 4 \end{array} \right) \end{array}$$

优化目标：

$$\text{minimize}_{\{b_u, b_i\}} \sum_{(u,i)} (r_{ui} - \hat{r}_{ui})^2$$

遍历训练集

- 利用训练集获得模型参数
- 利用测试集验证性能

$$\text{RMSE} = \sqrt{\sum_{(u,i)} \frac{(r_{ui} - \hat{r}_{ui})^2}{C}}.$$

Netflix电影个性化推荐

$$\|\mathbf{A}\mathbf{b}-\mathbf{c}\|_2^2$$

$$\text{minimize}_{\{b_u,b_i\}} \sum_{(u,i)} (r_{ui}-\hat{r}_{ui})^2$$

$$\hat{r}_{ui} = \bar{r} + b_u + b_i$$

30个方程，对应训练集中样本数

$$\begin{array}{cc} & \begin{array}{ccccccccc}1&2&3&\cdots&10&A&B&\cdots&E\end{array} \\ \begin{array}{c}1\\2\\\vdots\\30\end{array} & \begin{pmatrix}1&0&0&\cdots&0&1&0&\cdots&0\\0&0&1&\cdots&0&1&0&\cdots&0\\\vdots&&&&\ddots&&&\vdots&\\0&0&0&\cdots&1&0&0&\cdots&1\end{pmatrix}\end{array}\boxed{\begin{pmatrix}b_1\\b_2\\\vdots\\b_{10}\\b_A\\b_B\\\vdots\\b_E\end{pmatrix}}-\begin{pmatrix}r_{1A}-\bar{r}\\r_{3A}-\bar{r}\\\vdots\\r_{10E}-\bar{r}\end{pmatrix}$$

15个待估参数

$$(\mathbf{A}^T\mathbf{A})\mathbf{b}=\mathbf{A}^T\mathbf{c}$$

（最小二乘解）

Netflix电影个性化推荐

$$\mathrm{b}_u^* = [0.62, 0.42, -0.28, -1.78, 0.52, 0.49, -1.24, 0.45, 0.40, 0.23]^T$$

$$\mathrm{b}_i^* = [0.72, -1.20, 0.60, -0.60, 0.33]^T$$

$$\hat{\mathrm{R}} = \begin{array}{cc} & \begin{array}{ccccc} A & B & C & D & E \end{array} \\ \begin{array}{c} 1 \\ 2 \\ 3 \\ 4 \\ 5 \\ 6 \\ 7 \\ 8 \\ 9 \\ 10 \end{array} & \begin{pmatrix} 5.00 & 3.09 & 4.90 & - & \boxed{4.62} \\ - & 2.89 & 4.69 & \boxed{3.49} & 4.42 \\ 4.10 & 2.19 & - & \boxed{2.78} & 3.71 \\ - & \boxed{1.00} & 2.49 & 1.29 & 2.22 \\ 4.90 & - & \boxed{4.79} & 3.58 & 4.51 \\ \boxed{4.88} & 2.96 & - & 3.56 & 4.48 \\ 3.15 & \boxed{1.23} & 3.03 & 1.82 & - \\ 4.84 & \boxed{2.92} & 4.72 & - & 4.44 \\ \boxed{4.84} & 2.92 & 4.72 & 3.51 & - \\ \boxed{4.61} & - & 4.49 & 3.29 & 4.22 \end{pmatrix} \end{array}$$

RMSE=

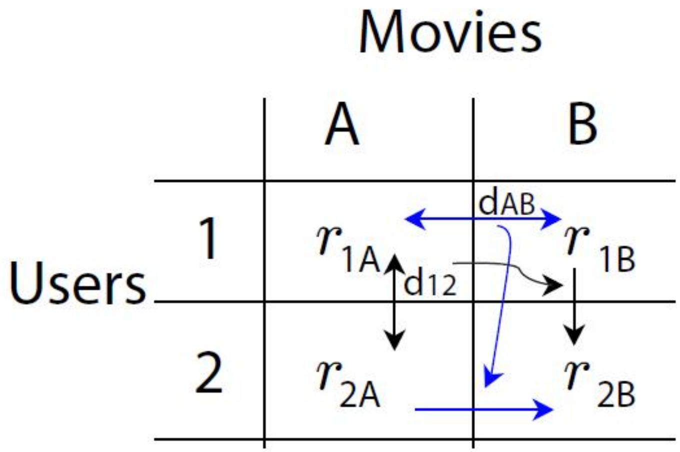
0.51 (训练集)

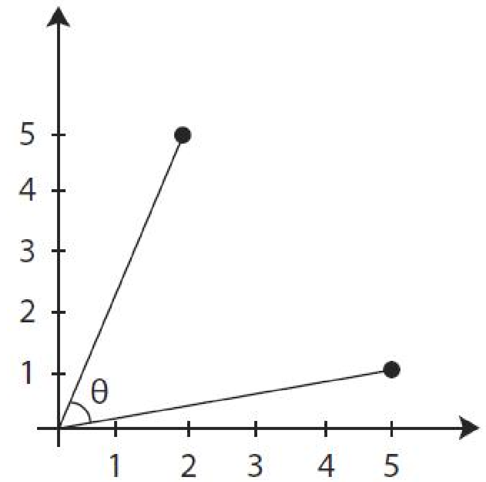
0.58 (测试集)

Netflix电影个性化推荐——实现问题

- 考虑相关性的加权预测

1.用户间的相关性：用户品味
2.电影的质量





- 时变因素

$$b_i(t) = b_i + b_{i,\text{bin}(t)} \quad b_u(t) = b_u + \sigma_u(t) + b_{u,t}$$

$$\hat{r}_{ui}(t) = \bar{r} + b_i(t) + b_u(t)$$

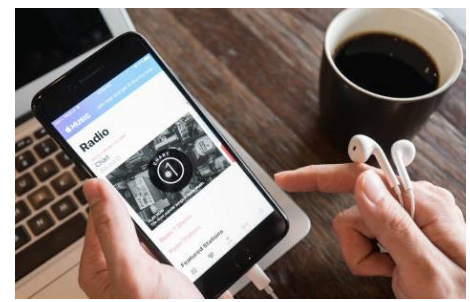
内容提要

- 最小二乘法
- 应用举例1：Netflix电影推荐
- 应用举例2：信道估计

无线通信中的信道估计

$y=hx+n$

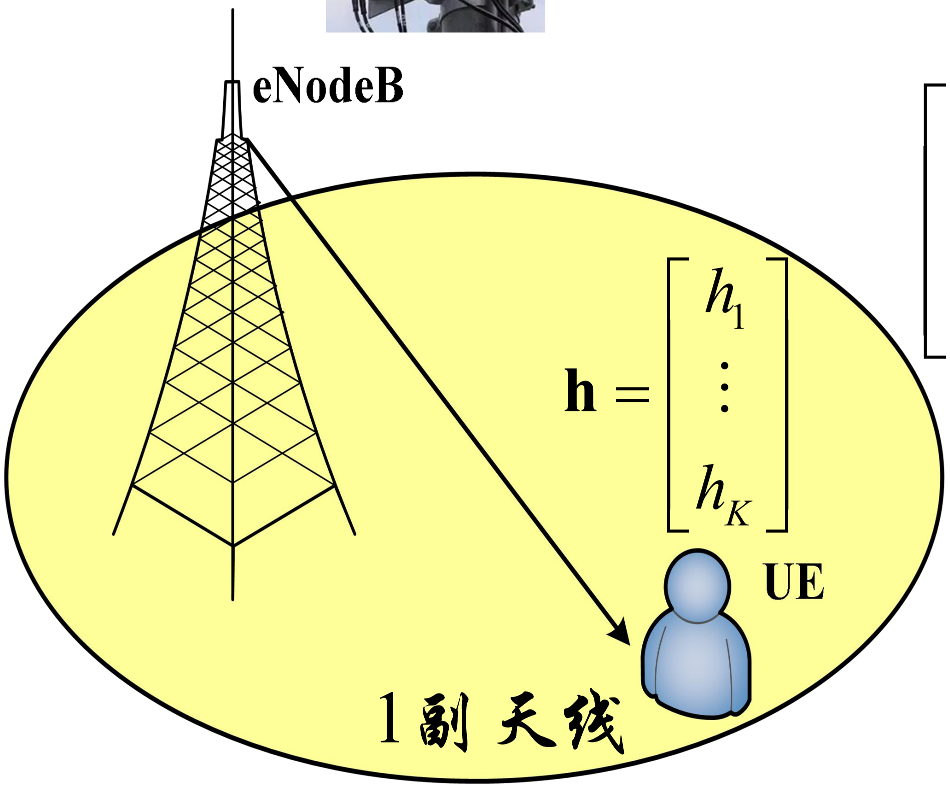
$y=h^{T}x+n$



- 信道估计是接收机实现相干检测的基础
- 信道估计是无线通信系统的关键技术之一

基于导频的LS信道估计

K副天线



导频	数 据	导频	数 据

时刻t

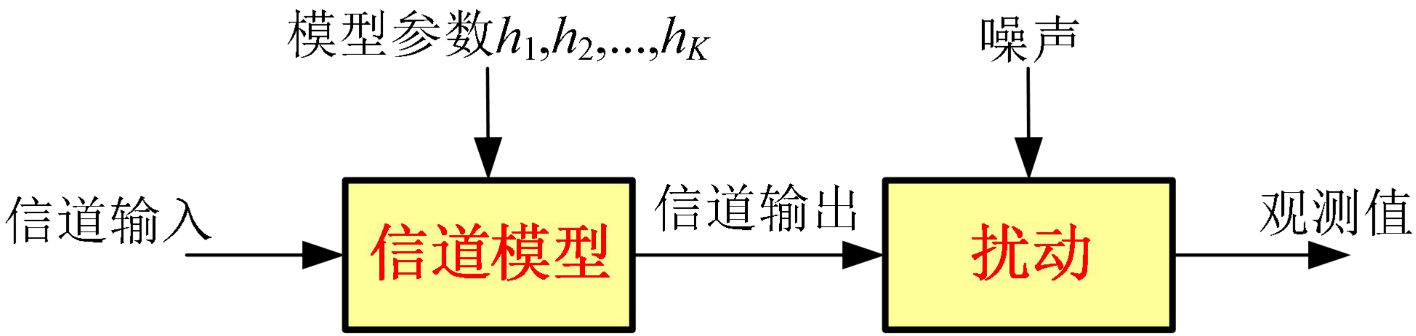
$$\begin{bmatrix} y_1 \\ \vdots \\ y_N \end{bmatrix} = \begin{bmatrix} x_{11} & x_{12} & \cdots & x_{1K} \\ x_{21} & x_{22} & \cdots & x_{2K} \\ & \vdots & & \\ x_{N1} & x_{N2} & \cdots & x_{NK} \end{bmatrix} \begin{bmatrix} h_1 \\ \vdots \\ h_K \end{bmatrix} + \begin{bmatrix} n_1 \\ \vdots \\ n_N \end{bmatrix}$$

天线1的导频序列

用于估计h_1的导频序列

信道估计本质上是系统辨识问题

基于导频的LS信道估计



$$\begin{bmatrix} y_1 \\ \vdots \\ y_N \end{bmatrix} = \begin{bmatrix} x_{11} & x_{12} & \cdots & x_{1K} \\ x_{21} & x_{22} & \cdots & x_{2K} \\ & \vdots & & \\ x_{N1} & x_{N2} & \cdots & x_{NK} \end{bmatrix} \begin{bmatrix} h_1 \\ \vdots \\ h_K \end{bmatrix} + \begin{bmatrix} n_1 \\ \vdots \\ n_N \end{bmatrix}$$

寻找能够对观测值给出最佳解释的模型参数

$$\text{minimize } \|\mathbf{y} - \mathbf{X}\mathbf{h}\|^2 \Rightarrow \hat{\mathbf{h}} = \left(\mathbf{X}^T\mathbf{X}\right)^{-1}\mathbf{X}^T\mathbf{y}$$

基于导频的LS信道估计

$$\begin{bmatrix} y_1 \\ \vdots \\ y_N \end{bmatrix} = \begin{bmatrix} x_{11} & x_{12} & \cdots & x_{1K} \\ x_{21} & x_{22} & \cdots & x_{2K} \\ & & \vdots & \\ x_{N1} & x_{N2} & \cdots & x_{NK} \end{bmatrix} \begin{bmatrix} h_1 \\ \vdots \\ h_K \end{bmatrix} + \begin{bmatrix} n_1 \\ \vdots \\ n_N \end{bmatrix}$$

$\mathbf{x}_1$ $\mathbf{x}_K$

1. 保证可逆？
2. 什么样的导频序列是最好？

$$\text{minimize } \|\mathbf{y} - \mathbf{X}\mathbf{h}\|^2 \Rightarrow \hat{\mathbf{h}} = \left(\mathbf{X}^T\mathbf{X}\right)^{-1}\mathbf{X}^T\mathbf{y}$$

$$\left(\mathbf{X}^T\mathbf{X}\right) = \begin{bmatrix} \mathbf{x}_1^T \\ \vdots \\ \mathbf{x}_K^T \end{bmatrix} \begin{bmatrix} \mathbf{x}_1 & \cdots & \mathbf{x}_K \end{bmatrix} = \begin{bmatrix} \|\mathbf{x}_1\|^2 & 0 & \cdots & 0 \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \cdots & \|\mathbf{x}_K\|^2 \end{bmatrix} \xrightarrow{\text{能量限制为1}} \hat{\mathbf{h}} = \mathbf{X}^T\mathbf{y} = \begin{bmatrix} \mathbf{x}_1^T \\ \vdots \\ \mathbf{x}_K^T \end{bmatrix}\mathbf{y} = \begin{bmatrix} \mathbf{x}_1^T\mathbf{y} \\ \vdots \\ \mathbf{x}_K^T\mathbf{y} \end{bmatrix}$$

导频正交时各个信道估计可以独立进行

谢谢大家！