http://staff.ustc.edu.cn/~ynyang/vector

第二十一讲 分类预测

2025.6.4

机器学习的6个核心算法(吴恩达,Andrew Ng)

- Linear Regression: Straight & Narrow
- Logistic Regression: Follow the Curve
- Gradient Descent: It's All Downhill
- Neural Networks: Find the Function
- Decision Trees: From Root to Leaves
- K-Means Clustering: Group Think

分类（预测、判别）

预测(prediction)：对未知的随机变量进行“估计”。当待预测随机变量是类别的时候，预测也称为分类(classification)或判别(discriminant)（当待预测随机变量是连续变量的时候，称为回归）。

数据：$(y_i, \mathbf{x}_i), i = 1, \ldots, n,$
求解 $p(\mathbf{x}, \theta) = E(y|\mathbf{x})$

y：类别，比如y取0-1类别

$\mathbf{x}$：自变量或特征feature

$E(y|\mathbf{x}) = p(y = 1|\mathbf{x})$是概率

训练

参数估计$\hat{\theta}$，拟合曲线$y = p(\mathbf{x}, \hat{\theta})$

$\hat{\theta}$=argmin$(-\sum[y_i \log p_i + (1 - y_i)\log(1 - p_i)])$

$p_i = p(\mathbf{x}_i, \theta) = P(y_i = 1|\mathbf{x}_i, \theta)$

回归: $\hat{\theta}$=argmin$\sum(y_i - f(\mathbf{x}_i, \theta))^2$

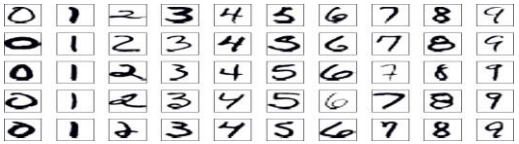
判别
预测

$\hat{y} = p(\mathbf{x}, \hat{\theta})$

预测

模式识别

例1 (手写体识别). $n = 50$ 个数字0−9手写体样本如右图。每个手写数字是16×16像素图像，每个像素点的值1(黑)或0(白). 将像素强度矩阵拉直成R^p向量, $p = 196$。



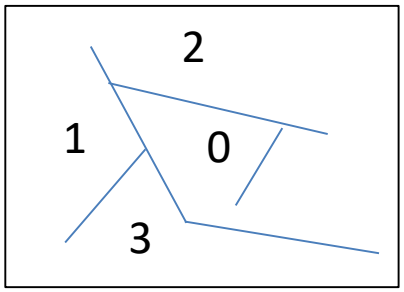
数据:

$\mathbf{x}_1, \ldots, \mathbf{x}_n \in R^p$: 手写数字的像素向量,

$y_1, \ldots, y_n$: 手写体的真实标签 (0−9, 类)。

训练预测(判别)准则:

将R^p划分成10个区域，与0−9对应(下图)



直线划分：线性判别、(线性)logistic回归

如果划分准则是曲线，则是非线性预测，比如神经网络。

预测:

判别新的手写体  是什么数字

⇔ 其向量表示 $\mathbf{x}$ 落在上图哪个区域?

两个学派

统计学分为两大流派：频率学派（Fisher学派，古典）和贝叶斯学派（条件概率），随着人工智能的发展，贝叶斯学派越来越被重视。

对于判别分析，Fisher的方法称为Fisher线性判别分析(LDA: linear discriminant analysis)，贝叶斯方法可得到类似的线性判别以及二次判别或其它非线性判别.

Ronald Fisher （1890-1962）英国统计学家.

组间与组内平方和: $T = W + B$

Thomas Bayes （1701-1761），英国统计学家、哲学家，发现了贝叶斯公式。

编码与解码：$p(\mathbf{z}|\mathbf{x})$与 $p(\mathbf{x}|\mathbf{z}), \mathbf{z}: latent$

$$p(y|\mathbf{x}) = \frac{P(\mathbf{x}|y)P(y)}{\sum_i P(\mathbf{x}|y=i)P(y=i)},\ y: \text{类别}$$

贝叶斯判别与logistic回归

两类判别/预测

假设$\mathbf{x} \in R^p$, 类别两类标号$y = 0,1$, 假设两个类的概率密度

$$\mathbf{x}|_{y=1} \sim f_1,\ \mathbf{x}|_{y=0} \sim f_0$$

如果f_1，f_0已知 (从数据$(y_i, \mathbf{x}_i)$, $i = 1, \ldots, n$训练/估计得到)，我们希望判别/预测$\mathbf{x}$所属类别（来自于f_1还是f_0？）

贝叶斯判别

一个自然的分类方式是比较概率，比如：

贝叶斯分类判别(分类)：

若 $P(y = 1|\mathbf{x}) > c$, 则预测$y = 1$.

阈值 c 是常数,不同地方出现的 c 未必相同。

记第一类在总体中的比例$p = P(y = 1)$, 利用贝叶斯公式得

$$P(y = 1|\mathbf{x}) = \frac{pf_1(\mathbf{x})}{pf_1(\mathbf{x}) + (1-p)f_0(\mathbf{x})}$$

$$= \frac{pf_1(\mathbf{x})/(1-p)f_0(\mathbf{x})}{1 + pf_1(\mathbf{x})/(1-p)f_0(\mathbf{x})} \triangleq \frac{\exp(a + b(\mathbf{x}))}{1 + \exp(a + b(\mathbf{x}))}.$$

其中 $b(\mathbf{x}) = \log\left(\frac{f_1(\mathbf{x})}{f_0(\mathbf{x})}\right), a = \log\left(\frac{p}{1-p}\right), p = P(y = 1).$

Logistic回归

因此，利用贝叶斯公式我们得到回归函数 $\mathrm{E}(y|\mathbf{x}) = P(y=1|\mathbf{x})$ 的logistic回归的一般形式：

(一般)logistic回归：$P(y=1|\mathbf{x}) = \dfrac{\exp(a+b(\mathbf{x}))}{1+\exp(a+b(\mathbf{x}))}$,

其中 $b(\mathbf{x}) = \log\left(\dfrac{f_1(\mathbf{x})}{f_0(\mathbf{x})}\right), a = \log\left(\dfrac{p}{1-p}\right), p = P(y=1).$

特例1：线性判别

如果两类分布是等方差正态：$f_1 = N_p(\boldsymbol{\mu}_1, \Sigma), f_0 = N_p(\boldsymbol{\mu}_0, \Sigma)$， 则

$$b(\mathbf{x}) = \log\left(\frac{f_1(\mathbf{x})}{f_0(\mathbf{x})}\right) = (\boldsymbol{\mu}_1 - \boldsymbol{\mu}_0)^{\mathsf{T}}\Sigma^{-1}\mathbf{x} - \frac{1}{2}(\boldsymbol{\mu}_1^{\mathsf{T}}\Sigma^{-1}\boldsymbol{\mu}_1 - \boldsymbol{\mu}_0^{\mathsf{T}}\Sigma^{-1}\boldsymbol{\mu}_0),$$

是$\mathbf{x}$的线性函数，此时得到通常的（线性）logistic回归模型：

(线性)logistic回归：$P(y=1|\mathbf{x}) = \dfrac{\exp(\alpha+\boldsymbol{\beta}^{\mathsf{T}}\mathbf{x})}{1+\exp(\alpha+\boldsymbol{\beta}^{\mathsf{T}}\mathbf{x})}$

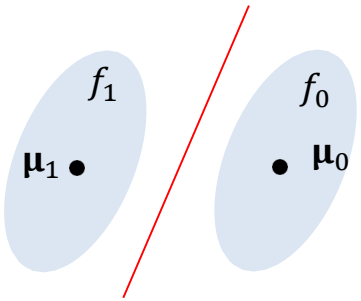
其中 $\boldsymbol{\beta}^{\mathsf{T}} = (\boldsymbol{\mu}_1 - \boldsymbol{\mu}_0)^{\mathsf{T}}\Sigma^{-1}\mathbf{x},\ \alpha = a - \frac{1}{2}(\boldsymbol{\mu}_1^{\mathsf{T}}\Sigma^{-1}\boldsymbol{\mu}_1 - \boldsymbol{\mu}_0^{\mathsf{T}}\Sigma^{-1}\boldsymbol{\mu}_0)$

此时， 贝叶斯判别或logistic回归判别是经典的Fisher线性判别（LDA：linear discriminant analysis)：
如果$P(y=1|\mathbf{x})>c$，等价地若

$$(\boldsymbol{\mu}_1-\boldsymbol{\mu}_0)^\top\Sigma^{-1}\left(\mathbf{x}-\frac{\boldsymbol{\mu}_1+\boldsymbol{\mu}_0}{2}\right)>c,$$

则预测$y=1$。

LDA经典的Fisher投影求法参见下页。



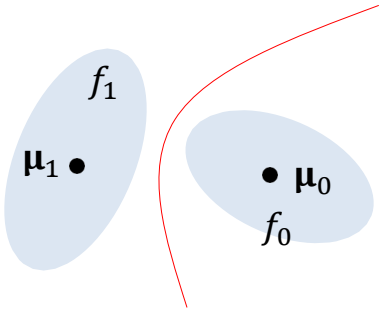
线性分类

特例2：二次判别

若$f_1=N_p(\boldsymbol{\mu}_1,\Sigma_1)$, $f_0=N_p(\boldsymbol{\mu}_0,\Sigma_0)$，$\Sigma_1\neq\Sigma_0$，则$b(\mathbf{x})=\log\left(\frac{f_1(\mathbf{x})}{f_0(\mathbf{x})}\right)=\boldsymbol{\beta}^\top\mathbf{x}+\mathbf{x}^\top Q\mathbf{x}=\mathbf{x}$的二次函数，此时

$$P(y=1|\mathbf{x})=\frac{\exp(\alpha+\boldsymbol{\beta}^\top\mathbf{x}+\mathbf{x}^\top Q\mathbf{x})}{1+\exp(\alpha+\boldsymbol{\beta}^\top\mathbf{x}+\mathbf{x}^\top Q\mathbf{x})}$$

判别准则：如果$P(y=1|\mathbf{x})>c$,
等价地若$\boldsymbol{\beta}^\top\mathbf{x}+\boldsymbol{\gamma}^\top Q\boldsymbol{\gamma}>c$, 则预测$y=1$。



二次分类

神经网络方法容许$b(\mathbf{x})$具有一般的非线性形式，即非线性分类或判别。

线性判别(LDA，linear discriminant analysis)

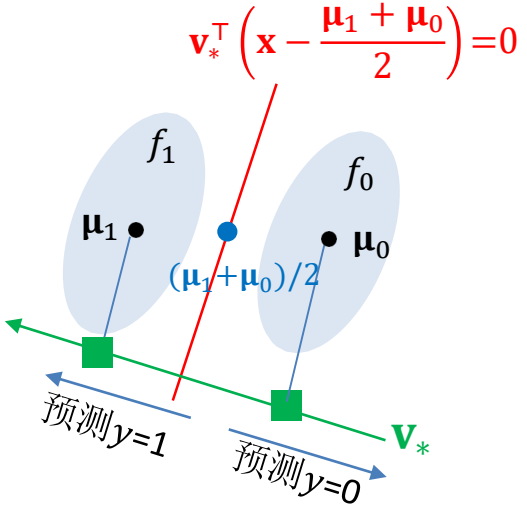
Fisher投影方法:

假设$f_1 = N_p(\boldsymbol{\mu}_1, \Sigma), f_0 = N_p(\boldsymbol{\mu}_0, \Sigma)$，求$\mathbf{v}$使得两类数据在$\mathbf{v}$方向区分度最大:

$$\max_{\mathbf{v}} \frac{(\mathbf{v}^\top \boldsymbol{\mu}_1 - \mathbf{v}^\top \boldsymbol{\mu}_0)^2}{\mathbf{v}^\top \Sigma \mathbf{v}}$$

$\Rightarrow$ 最优方向$\mathbf{v}_* = \Sigma^{-1}(\boldsymbol{\mu}_1 - \boldsymbol{\mu}_0)$

(区分度与组间方差/组内方差之比BW^{-1}有关)



Fisher线性判别准则:

若 $\mathbf{v}_*^\top\left(\mathbf{x}-\frac{\boldsymbol{\mu}_1+\boldsymbol{\mu}_0}{2}\right) > c$，则判别/预测$y = 1$

假设两类数据
$\mathbf{x}_{11}, \dots, \mathbf{x}_{1n_1} \sim f_1$;
$\mathbf{x}_{01}, \dots, \mathbf{x}_{0n_0} \sim f_0$。
判别准则中$\boldsymbol{\mu}_1, \boldsymbol{\mu}_0, \Sigma$ 以$\bar{\mathbf{x}}_1, \bar{\mathbf{x}}_0, S$ 代入

注: 当$c = 0$, 边界

$\mathbf{v}_*^\top\left(\mathbf{x}-\frac{\boldsymbol{\mu}_1+\boldsymbol{\mu}_0}{2}\right)=0 \Leftrightarrow f_1(\mathbf{x}) = f_0(\mathbf{x})$ 等概线

$\Leftrightarrow (\mathbf{x}-\boldsymbol{\mu}_1)^\top \Sigma^{-1}(\mathbf{x}-\boldsymbol{\mu}_1) < (\mathbf{x}-\boldsymbol{\mu}_0)^\top \Sigma^{-1}(\mathbf{x}-\boldsymbol{\mu}_0)$，等(马氏)距线

多类判别和多项回归

假设$\mathbf{x} \in R^p$, 类别标号$y = 0,1,\ldots,K-1$, 假设各类的概率密度

$$\mathbf{x}|_{y=k} \sim f_k,\ 0,1,\ldots,K-1$$

记$p_k = P(y=k)$, 记 $a_k = \log\left(\frac{p_k}{p_0}\right)$, $b_k(\mathbf{x}) = \log\left(\frac{f_k(\mathbf{x})}{f_0(\mathbf{x})}\right)$, 则有多项回归(multinomial regression).

$$P(y=k|\mathbf{x}) = \frac{P(\mathbf{x}|y=k)p_k}{P(\mathbf{x})} = \frac{f_k(\mathbf{x})p_k}{\sum_{i=1}^{K} f_i(\mathbf{x})p_i}$$

$$= \frac{\exp(a_k + b_k(\mathbf{x}))}{\sum_{i=1}^{K} \exp(a_i + b_i(\mathbf{x}))},\ a_0 = 0,\ b_0(\mathbf{x}) = 0$$

特例: 多类Fisher线性判别

若$f_k = N_p(\boldsymbol{\mu}_k, \Sigma)$, 则模型具有普通形式（指数上线性）：

$$P(y=k|\mathbf{x}) = \frac{\exp(\alpha_k + \boldsymbol{\beta}_k^{\top}\mathbf{x})}{\sum_{i=1}^{K} \exp(\alpha_i + \boldsymbol{\beta}_k^{\top}\mathbf{x})}$$

Fisher多类线性判别准则:

若 $k_0 = \arg\max_{k=0,..,.K-1} P(y=k|\mathbf{x})$, 判别$y = k_0$.

$b_k(\mathbf{x})$非线性?

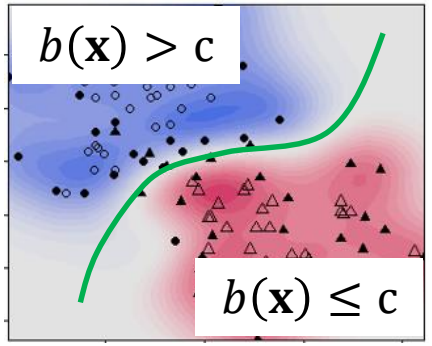
神经网络

以二类预测为例（多类问题类似）。

神经网络假设 logistic回归中$b(\mathbf{x})$非线性

$$P(y=1|\mathbf{x})=\frac{\exp\big(a+b(\mathbf{x})\big)}{1+\exp\big(a+b(\mathbf{x})\big)}$$

如果$P(y=1|\mathbf{x})>c \Leftrightarrow b(\mathbf{x})>c$（右图），预测$y=1$。

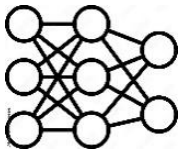


如何生成非线性函数$b(\mathbf{x})$？多个ReLU(或其它非线性激活)的组合：

ReLU(分段线性): $x_+=\begin{cases}x, x>0\\0, x\leq 0\end{cases}$

$$b(x)=\beta_1(x-b_1)_+ +\beta_2(x-b_2)_+ =\begin{cases}0, & x\leq b_1\\ \beta_1(x-b_1), & b_1<x\leq b_2\\ \beta_1(x-b_1)+\beta_2(x-b_2), & x>b_2\end{cases}$$

深度神经网络 DNN



假设logistic回归

$$P(y=1|\mathbf{x}) = \frac{\exp\big(a+b(\mathbf{x})\big)}{1+\exp\big(a+b(\mathbf{x})\big)}$$

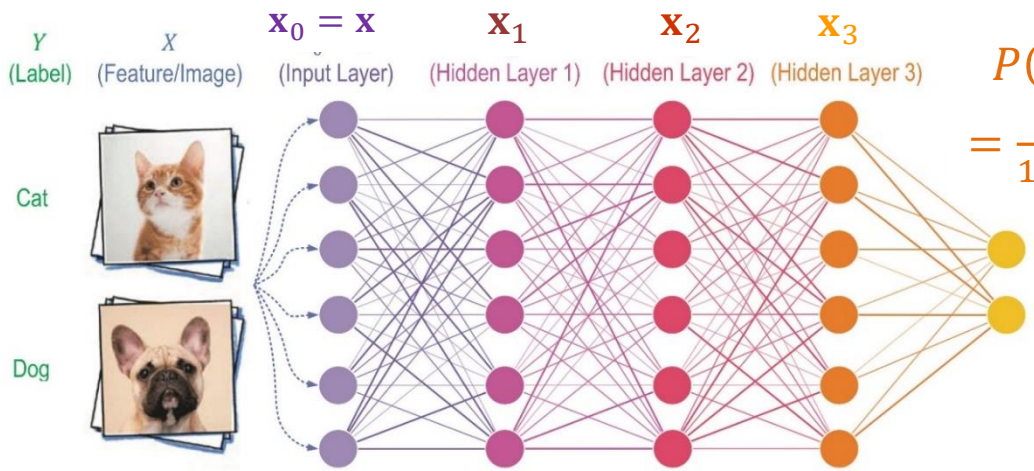
深度神经网络(Deep neural network, DNN) 使用若干次仿射变换+正部激活递归复合而成$b(\mathbf{x})$:

$$b(\mathbf{x}) = \boldsymbol{\beta}^{\mathsf{T}}\mathbf{x}_h = \boldsymbol{\beta}^{\mathsf{T}} f^h\Big(f^{h-1}(\cdots f^1(\mathbf{x}))\Big)$$

其中递归变换:

$$\mathbf{x}_0 = \mathbf{x},\ \ \mathbf{x}_k = f^k(\mathbf{x}_{k-1}) = [A_k\mathbf{x}_{k-1} + \mathbf{b}_k]_+,\ \ k = 1,\dots,h$$

其中A_k是矩阵参数, $\mathbf{b}_k$是偏置向量参数(bias).



训练/拟合

假设数据为$(y_i, \mathbf{x}_i), i = 1, \ldots, n, y_i = 0,1, \mathbf{x}_i \in R^p$, 神经网络假设

$$p_i(\boldsymbol{\theta}) = P(y_i = 1|\mathbf{x}_i) = \frac{\exp\big(a + b(\mathbf{x}_i)\big)}{1 + \exp\big(a + b(\mathbf{x}_i)\big)}$$

其中$b(\mathbf{x}_i) = \boldsymbol{\beta}^{\top} f^h\Big(f^{h-1}(\cdots f^1(\mathbf{x}_i))\Big)$, $f^k(\boldsymbol{z}) = [A_k\boldsymbol{z} + \mathbf{b}_k]_+$, 似然

$$L(\boldsymbol{\theta}) = \prod_{i=1}^{n} p_i(\boldsymbol{\theta})^{y_i}(1 - p_i(\boldsymbol{\theta}))^{1-y_i}$$

$$= \prod_{i=1}^{n} \left(\frac{\exp\big(a + b(\mathbf{x}_i)\big)}{1 + \exp\big(a + b(\mathbf{x}_i)\big)}\right)^{y_i} \left(\frac{1}{1 + \exp\big(a + b(\mathbf{x}_i)\big)}\right)^{1-y_i}$$

极小化交叉熵损失$(-logL)$

$$J(\boldsymbol{\theta}) = -\sum_{i=1}^{n} y_i \log\big(p_i(\boldsymbol{\theta})\big) + (1 - y_i)\log\big(1 - p_i(\boldsymbol{\theta})\big)$$

梯度下降法求解$\boldsymbol{\theta}$:

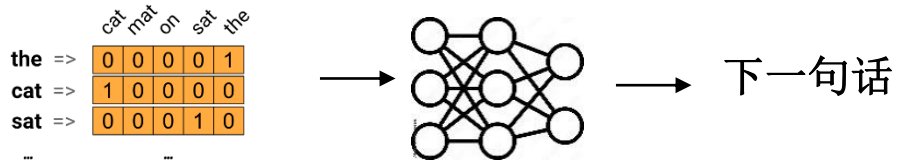
$$\boldsymbol{\theta}^{(k)} = \boldsymbol{\theta}^{(k-1)} - \eta \partial J/\partial \boldsymbol{\theta}|_{\boldsymbol{\theta}=\boldsymbol{\theta}^{(k-1)}},$$

η: learning rate

具体的embedding/encoding layer →深度神经网络(DNN)→模型结构的改变→各种深度学习模型。其中我们所学到的各种embedding、模型结构都可能与DNN建立联系。例如：

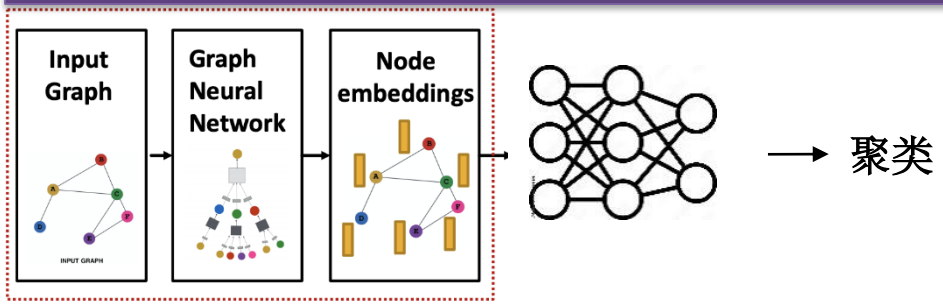
文本模型

文本one-hot表示　　DNN　　预测



图神经网络 (GNN)

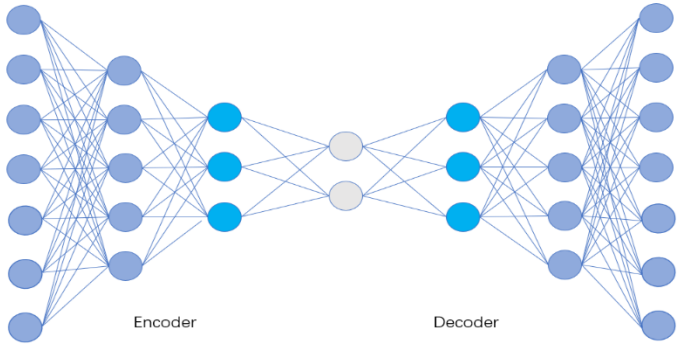
图特征向量　　DNN　　无监督学习



自编码器 Autoencoder

编码 $p(\mathbf{z}|\mathbf{x})$　　少量潜变量$\mathbf{z}$　　解码$p(\mathbf{x}|\mathbf{z})$

潜变量: 主成分、潜因子



正态假设下(13讲P36)

- 解码$p(\mathbf{x}|\mathbf{z}) \Leftrightarrow \mathbf{x} = L\mathbf{z} + \boldsymbol{\varepsilon}$
  $p(\mathbf{x}|\mathbf{z})$: $N_p(0, LL^{\mathsf{T}} + \Psi)$
- 编码$p(\mathbf{z}|\mathbf{x})$ ⇔预测因子(score)
  $p(\mathbf{z}|\mathbf{x})$: $N_m(L^{\mathsf{T}}(LL^{\mathsf{T}} + \Psi)^{-1}\mathbf{x}, I_m - L^{\mathsf{T}}(LL^{\mathsf{T}} + \Psi)^{-1}L)$

阈值 c 如何确定？分类效果如何评价？

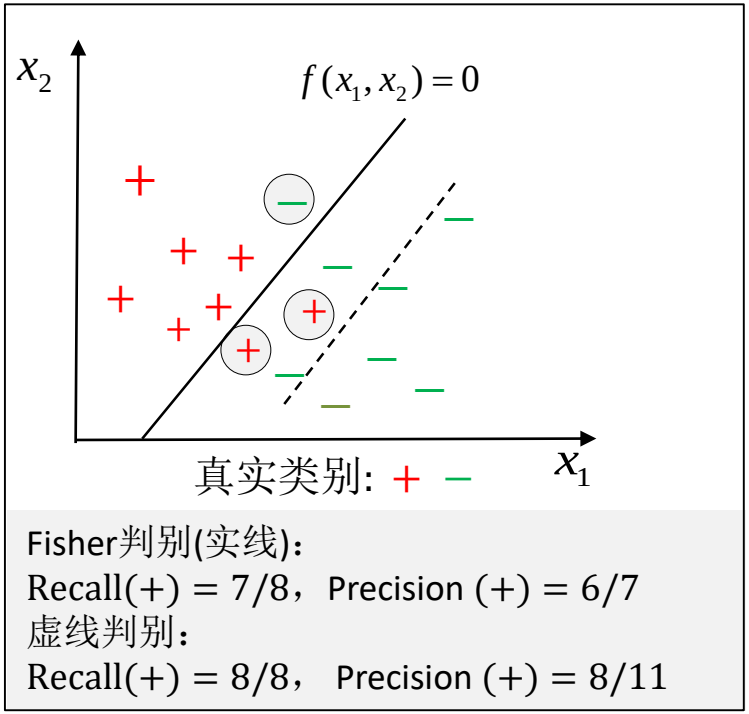
评价判别效果

以Fisher 线性判别(LDA)为例

$$f(x_1, x_2) = ax_2 + bx_1 + c$$

若$f(x_1, x_2) > 0$，判别为阳性 +；

否则，判为阴性 − .



Fisher判别(实线):

Recall(+) = 7/8，Precision (+) = 6/7

虚线判别:

Recall(+) = 8/8， Precision (+) = 8/11

Fisher判别将2个+ 错判为−；1个−错判为+。

如果正确判定+是重要的（比如疾病），那么应该下调判别阈值，比如： $f(x_1, x_2) > -1$ 时，判为+，这可以增加正确判别+的概率：

$$\text{Recall}(+) = P(\text{判为} + \,|\text{真实为}+),$$

极端地，若判别阈值为负无穷大，则所有点判为+ ，recall=1，但此时所有判别为+的对象中有很多误判，精度较小

$$\text{Precision}(+) = \frac{\text{正确判别为+的个数}}{\text{所有判别为+的个数}}$$

若关注重点是 - ，同样定义Recall(-),Precision(-). 无论如何，评估分类效果需综合考虑准确度和精确度，比如两者的几何平均F1。

判别效果度量

测试数据与训练数据格式完全相同（含真实的类别标号），但不参与训练。用训练得到的方法预测测试数据的类别标号，与真实类别标号比较考察效果。

两个类：＋positive阳性; – negative阴性，
四种结果：TP: true positive； FP: false positive； FN: false negative TN: true negative

		真正的类别 +	真正的类别 –
判定的类别	+	TP	FP
	–	FN	TN

常用的分类正确率如下（前4个2 × 2的边际概率）：

准则	定义	解释
召回率recall(+), 灵敏度Sensitivity	TP/(TP+FN)	真阳性被判对的比例
精确度Precision (+)	TP/(TP+FP)	判为阳性的判别中正确的比例
准确度 accuracy	(TP+TN)/(TP+FP+TN+FN)	所有判别中正确判别的比例
F1 score=2/(1/recall+1/precision)	2TP/(2TP+FP+FN)	灵敏度和精确度的几何平均

召回率(–)，特异度Specificity	**TN/(FP+TN)**	**真阴性被判对的比例**
精确度 Precision(–)	TN/(FN+TN)	判为阴性的判别中判对的比例

复习

课程主要内容

- ❖ 第1-10讲： 经典统计（normal-多元正态， 微积分）
- ❖ 第11-21讲：机器学习（singular-SVD， 线性代数）

- 8次纸质作业 （hw8不交)
- 4次上机作业（lab4不交）。

第1-10讲 (多元正态)

多元线性模型

假设线性结构 $\boldsymbol{\mu} = B^{\top}\mathbf{z}$

X 的第j列
$\mathbf{x}_{(j)} \sim N_n(\mathbf{0}, \sigma_{jj} I_n)$
球对称分布

$\boldsymbol{\mu} = \mathbf{0}, HX \overset{d}{=} X, H \in \mathcal{O}(n)$

一般球对称分布刻画：
$\mathbf{x}$~球对称分布$\Leftrightarrow \mathbf{x} \overset{d}{=} \mathbf{u}r$,
$\mathbf{u} \sim U(S^{n-1})$与$r$独立

$\mathbf{x}_{(j)} \overset{d}{=} \mathbf{u}_j r_j, \mathbf{u}_j$与$r_j$独立
$r_j^2 = \|\mathbf{x}_{(j)}\|^2 \sim \sigma_{jj}\chi_n^2$
$\mathbf{u}_j = \frac{\mathbf{x}_{(j)}}{\|\mathbf{x}_{(j)}\|} \sim U(S^{n-1})$

$\mathbf{x}_1, \ldots, \mathbf{x}_n$ iid $\sim N_p(\boldsymbol{\mu}, \Sigma)$,
矩阵 $X_{n\times p} = (\mathbf{x}_1, \ldots, \mathbf{x}_n)^{\top} = (\mathbf{x}_{(1)}, \ldots, \mathbf{x}_{(p)})$

$\boldsymbol{\mu} \neq \mathbf{0}$

MLE: $\hat{\boldsymbol{\mu}} = \bar{\mathbf{x}}, \hat{\Sigma} = S$

协方差结构约束(因子模型，假设检验)

$\boldsymbol{\mu} = \mathbf{0}, HX \overset{d}{=} X, H \in \mathcal{O}(n)$
X列向量球对称

注:右正交不变$XO \overset{d}{=} X, O \in \mathcal{O}(p)$
$\Leftrightarrow \Sigma = \sigma^2 I_p$ (X行向量球对称)

极大似然⇔极小化：
$\log|\Sigma(\boldsymbol{\theta})| + tr(\Sigma(\boldsymbol{\theta})^{-1}S)$

$X^{\top}X = \sum_{i=1}^{n} \mathbf{x}_i\mathbf{x}_i^{\top} \sim W_p(n, \Sigma)$

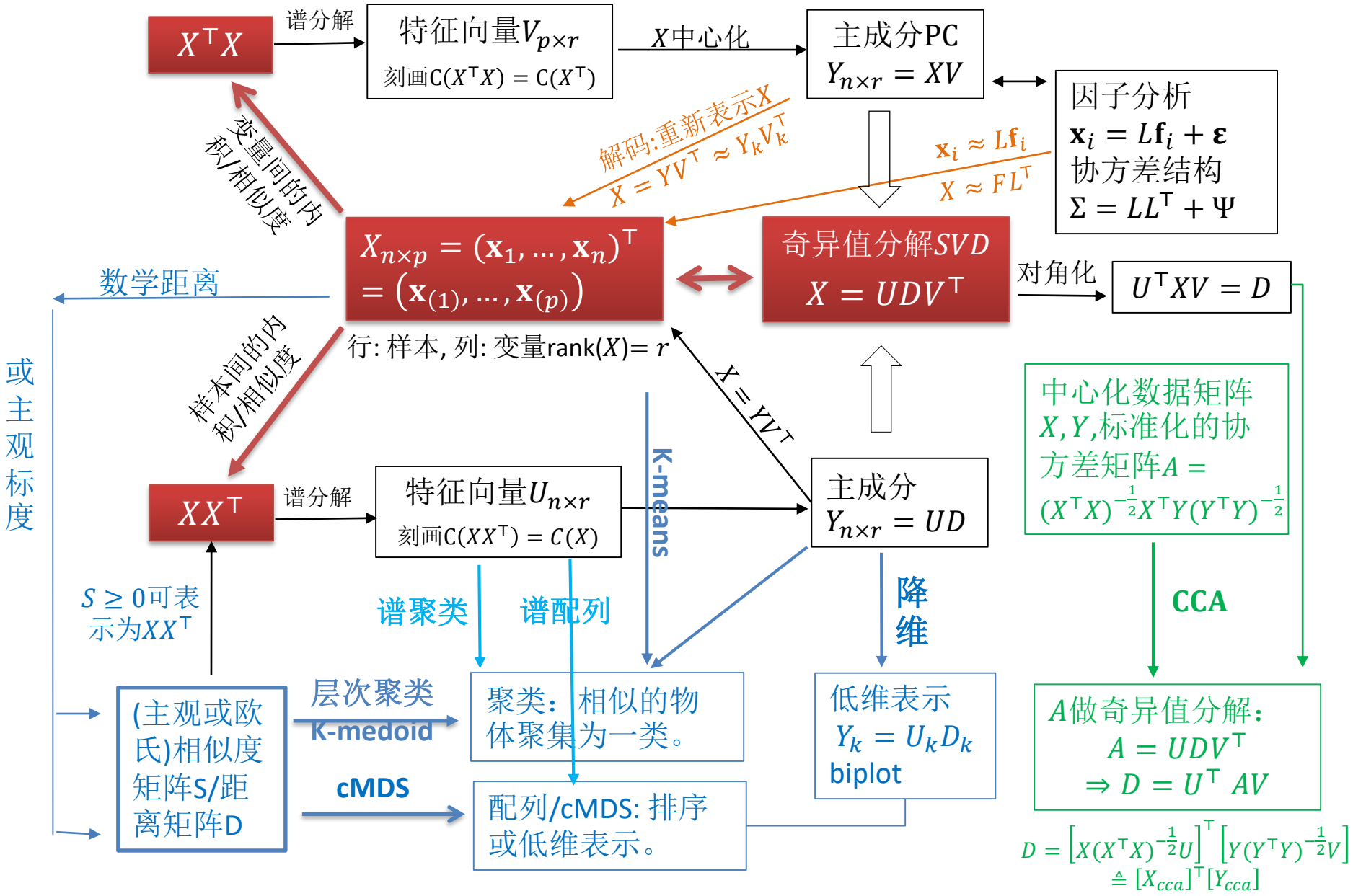
$X^{\top}X = (\mathbf{x}_{(i)}^{\top}\mathbf{x}_{(j)}) = (\mathbf{u}_i^{\top}\mathbf{u}_j r_i r_j)$
列向量(变量)之间的内积

关于正态分布的一般检验方法：
Wilks/似然比检验(近似检验)
$-2\log\left(\frac{\max\limits_{H0} L}{\max\limits_{H1} L}\right) = n\log\left(\frac{|\hat{\Sigma}_0|}{|S|}\right) \overset{H0}{\to} \chi_{df}^2$
df: H1,H0下参数个数之差。

两正态样本Hotelling T^2精确检验。原假设下服从cF 。

样本相关系数矩阵($\boldsymbol{\mu} = \mathbf{0}$)
$R = (\mathbf{u}_i^{\top}\mathbf{u}_j)/(n-1)$
其中 $\mathbf{u}_1, \ldots, \mathbf{u}_p \sim U(S^{n-1})$

第11-21讲（奇异值分解SVD）



考试范围

- 复习范围：课件+纸质作业 (不包含下面列表中删除内容及对应的作业题、不包含lab、不包含课件中的附录和虚线框).
- 考试内容：与纸质作业难度、类型类似 (除了hw5.4,5.5) 。
- 考试时间：2025-06-21 8:30-10:30；地点:2321. 带计算器。

注意考试地点有变更

课件	删除内容
第一讲：多元分析简介	
第二讲：球对称分布	
第三讲：球对称分布(II)	
第四讲：多元正态分布	
第五讲：高斯图模型	全部
第六讲：Wishart分布	P8-25(外微分)
第七讲：Wishart分布II	
第八讲：Wishart分布III	
第九讲：Hotelling's T^2 检验	
第十讲：多元线性回归模型	全部

	删除内容
第十一讲：主成分分析	
第十二讲：双标图	
第十三讲：因子分析	
第十四讲：结构方程模型	全部
第十五讲：奇异值分解	
第十六讲：典则相关分析	
第十七讲：列联表与对应分析	全部(18讲例1保留)
第十八讲：距离与相似系数	
第十九讲：多维标度法	
第二十讲：聚类分析	P36-46(谱分析)
第二十一讲：分类预测	P1-15(分类)

sample