第一章　绪论

1.1　概论

1.2　特征矢量和特征空间

1.3　随机矢量的描述

1.4　正态分布

1.2 特征矢量和特征空间

•特征矢量

一个分析对象的 n 个特征量测值分别为 $x_1, x_2, \cdots, x_n$ 它们构成一个 n 维特征矢量 $\boldsymbol{x}$， $\boldsymbol{x} = (x_1, x_2, \cdots, x_n)^{\mathrm{T}}$，$\boldsymbol{x}$ 是原对象（称为样本）的一种数学抽象，用来代表原对象，即为原对象的模式。

1.2　特征矢量和特征空间

•特征空间

　　对某对象的分类识别是对其模式，即它的特征矢量进行分类识别。

　　各种不同取值的x的全体构成了n维空间，这个n维空间称为特征空间，不同的场合特征空间可记为X^n、 ^{n}或Ω。特征矢量x便是特征空间中的一个点，所以特征矢量有时也称为特征点。

- 随机变量

由于量测系统随机因素的影响及同类不同对象的特征本身就是在特征空间散布的，同一个对象或同一类对象的某特征量测值是随机变量。由随机分量构成的矢量称为随机矢量。同一类对象的特征矢量在特征空间中是按某种统计规律随机散布的。

第一章　绪论

1.1　概论

1.2　特征矢量和特征空间

1.3　随机矢量的描述

1.4　正态分布

1. 3　　随机矢量的描述

(1) 随机矢量的分布函数

设 $\boldsymbol{X}=(X_1,X_2,\cdots,X_n)^{\mathrm{T}}$ 为随机矢量，$\boldsymbol{x}=(x_1,x_2,\cdots,x_n)^{\mathrm{T}}$ 为确定性矢量。 随机矢量 $\boldsymbol{X}$ 的联合概率分布函数定义为：

$$F(x_1,x_2,\cdots,x_n)=P(X_1\le x_1,X_2\le x_2,\cdots,X_n\le x_n)$$

写成矢量形式：

$$F(\boldsymbol{x})=P(\boldsymbol{X}\le\boldsymbol{x})$$

1.3 随机矢量的描述

➢联合概率密度函数

设 $\boldsymbol{X}=(X_1,X_2,\cdots,X_n)^{\mathrm{T}}$ 为随机矢量，$\boldsymbol{x}=(x_1,x_2,\cdots,x_n)^{\mathrm{T}}$ 为确定性矢量。随机矢量 $\boldsymbol{X}$ 的联合概率密度函数定义为：

$$p(x_1,x_2,\cdots,x_n)=\frac{\partial^n F(x_1,x_2,\cdots,x_n)}{\partial x_1 \partial x_2 \cdots \partial x_n}$$

1.3　随机矢量的描述

➢类概率密度函数

设集合由 c 类模式组成，第 i 类记为 ω_i，ω_i 类模式特征矢量有其自己的分布函数和密度函数。

ω_i 类的模式特征矢量的分布函数及密度函数分别定义为：

$$F(\mathbf{x}\mid\omega_i)=P(\mathbf{X}\le\mathbf{x}\mid\omega_i)$$

$$p(\mathbf{x}\mid\omega_i)=\partial^n F(x_1,x_2,\cdots,x_n\mid\omega_i)\big/\partial x_1\,\partial x_2\cdots\partial x_n$$

(2) 随机矢量的数字特征

➢ 均值矢量

n维随机矢量 X 的数学期望 $\boldsymbol{\mu}$ 定义为：

$$\boldsymbol{\mu} = \mathrm{E}[\boldsymbol{X}] = \bar{\boldsymbol{X}} = \begin{pmatrix} \mathrm{E}[X_1] \\ \mathrm{E}[X_2] \\ \cdots \\ \mathrm{E}[X_n] \end{pmatrix} = \int_{X^n} \boldsymbol{x} p(\boldsymbol{x}) \mathrm{d}\boldsymbol{x}$$

1.3 随机矢量的描述

➢ 条件期望

在模式识别中，经常以类别 ω_i 作为条件，在这种情况下随机矢量 X 的条件期望矢量定义为：

$$\boldsymbol{\mu}_i = \mathrm{E}[X \mid \omega_i] = \int_{X^n} \boldsymbol{x} p(\boldsymbol{x} \mid \omega_i)\mathrm{d}\boldsymbol{x}$$

1.3 随机矢量的描述

- 协方差矩阵

随机矢量$\boldsymbol{X}$的自协方差矩阵表征各分量围绕其均值的散布情况及各分量间的相关关系，其定义为：

$$\begin{aligned}\boldsymbol{\Sigma} &= \mathrm{E}[(\boldsymbol{X}-\bar{\boldsymbol{X}})(\boldsymbol{X}-\bar{\boldsymbol{X}})^{\mathrm{T}}] \\ &= \int_{X^n}(\boldsymbol{x}-\boldsymbol{\mu})(\boldsymbol{x}-\boldsymbol{\mu})^{\mathrm{T}}p(\boldsymbol{x})\mathrm{d}\boldsymbol{x} = (\sigma_{ij}^2)_{n\times n} \\ &= \mathrm{Cov}(\boldsymbol{X},\boldsymbol{X})\end{aligned}$$

➢ 自相关矩阵

随机矢量 X 的自相关矩阵定义为：

$$\boldsymbol{R}=\mathrm{E}\left[\boldsymbol{X}\boldsymbol{X}^{\mathrm{T}}\right]$$

由定义可知，X 的协方差矩阵和自相关矩阵间的关系是：

$$\boldsymbol{\Sigma}=\boldsymbol{R}-\bar{\boldsymbol{X}}\,\bar{\boldsymbol{X}}^{\mathrm{T}}=\boldsymbol{R}-\boldsymbol{\mu}\boldsymbol{\mu}^{\mathrm{T}}$$

1.3 随机矢量的描述

- 相关系数

$$r_{ij} = \sigma_{ij}^2 \big/ (\sigma_{ii}\,\sigma_{jj})$$

由布尼亚科夫斯基不等式知：

$$\left|\sigma_{ij}^2\right| \le \sigma_{ii}\sigma_{jj}$$

所以：

$$-1 \le r_{ij} \le 1$$

相关系数矩阵定义为：

$$\boldsymbol{R}_c = (r_{ij})_{n\times n}$$

➢ 协方差矩阵的非负定性

协方差矩阵和自相关矩阵都是对称矩阵。设$\boldsymbol{A}$为对称矩阵，对任意的矢量$\boldsymbol{x}$，$\boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x}$是$\boldsymbol{A}$的二次型。若对任意的$\boldsymbol{x}$恒有：

$$\boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x} \geq 0$$

则称$\boldsymbol{A}$为非负定矩阵。协方差矩阵是非负定的。

1. 3　随机矢量的描述

(3) 随机变量、随机矢量间的统计关系

➢ 不相关

1) 定义：随机矢量 $\boldsymbol{X}$ 的第个 i 分量 X_i 和第 j 个分量 X_j，若有 $\sigma_{ij}^2 = 0 (i \neq j)$，则称它们不相关。

2) 等价条件：X_i 和 X_j 不相关等价于：

$$\mathrm{E}[X_j X_j] = \mathrm{E}[X_i]\mathrm{E}[X_j]$$

3) 随机矢量 $\boldsymbol{X}$ 和 $\boldsymbol{Y}$ 不相关的充要条件是互协方差矩阵 $\mathrm{Cov}(\boldsymbol{X},\boldsymbol{Y}) = \boldsymbol{O}$，亦即 $\mathrm{E}[\boldsymbol{X}\boldsymbol{Y}^{\mathrm{T}}] = \mathrm{E}[\boldsymbol{X}]\mathrm{E}[\boldsymbol{Y}^{\mathrm{T}}]$

1.3　随机矢量的描述

➢ 正交

若随机矢量 X 和 Y 满足 $\mathrm{E}[X^{\mathrm{T}}Y]=0$，则称 X 和 Y 正交。

➢ 独立

随机矢量 X 和 Y 的联合概率密度函数 $p(x,y)$，若满足 $p(x,y)=p(x)p(y)$，则称 X 和 Y 独立。

➢ 两者关系

独立必不相关，反之不然。

(4) 随机矢量的变换

➢ 一般关系

设随机矢量$\boldsymbol{Y}$是另一随机矢量$\boldsymbol{X}$的连续函数，即：

$$\boldsymbol{Y}=\begin{pmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_n \end{pmatrix}=\begin{pmatrix} g_1(X_1,X_2,\cdots,X_n) \\ g_2(X_1,X_2,\cdots,X_n) \\ \vdots \\ g_n(X_1,X_2,\cdots,X_n) \end{pmatrix}\triangleq\begin{pmatrix} g_1(\boldsymbol{X}) \\ g_2(\boldsymbol{X}) \\ \vdots \\ g_n(\boldsymbol{X}) \end{pmatrix}\triangleq\boldsymbol{g}(\boldsymbol{X})$$

若它们的函数关系是一一对应的，则两个随机矢量的概率密度函数之间有关系：

$$p(\boldsymbol{y})=\frac{p(\boldsymbol{x})}{|J|}$$

J为雅可比行列式：

$$J=\begin{vmatrix} \dfrac{\partial g_1}{\partial x_1} & \dfrac{\partial g_1}{\partial x_2} & \cdots & \dfrac{\partial g_1}{\partial x_n} \\ \dfrac{\partial g_2}{\partial x_1} & \dfrac{\partial g_2}{\partial x_2} & \cdots & \dfrac{\partial g_2}{\partial x_n} \\ \vdots & \vdots & \vdots & \vdots \\ \dfrac{\partial g_n}{\partial x_1} & \dfrac{\partial g_n}{\partial x_2} & \cdots & \dfrac{\partial g_n}{\partial x_n} \end{vmatrix}$$

J表示变换后体积微元的变化，Y^n坐标系中体积微元 $\mathrm{d}y_1\mathrm{d}y_2\cdots\mathrm{d}y_n=|J|\mathrm{d}x_1\mathrm{d}x_2\cdots\mathrm{d}x_n$。

1.3 随机矢量的描述

$$\boldsymbol{Y}=\begin{pmatrix}Y_1\\Y_2\\\vdots\\Y_n\end{pmatrix}=\begin{pmatrix}a_{11}X_1+a_{12}X_2+\cdots+a_{1n}X_n\\a_{21}X_1+a_{22}X_2+\cdots+a_{2n}X_n\\\\a_{n1}X_1+a_{n2}X_2+\cdots+a_{nn}X_n\end{pmatrix}=\begin{pmatrix}a_{11}&a_{12}&\cdots&a_{1n}\\a_{21}&a_{22}&\cdots&a_{2n}\\\vdots&\vdots&\vdots&\vdots\\a_{n1}&a_{n2}&\cdots&a_{nn}\end{pmatrix}\begin{pmatrix}X_1\\X_2\\\vdots\\X_n\end{pmatrix}\triangleq\boldsymbol{AX}$$

此时，$J=|\boldsymbol{A}|$，则随机矢量$\boldsymbol{Y}$的概率密度函数

$$p(y)=p(x)/\|\boldsymbol{A}\|$$

➢ 数字特征关系

设 X 的均值矢量为μ_x, 协方差矩阵为Σ_x , 则 Y 的均值矢量

$$\mu_y = \mathrm{E}\left[Y\right] = A\mathrm{E}\left[X\right] = A\mu_x$$

Y 的协方差阵

$$\Sigma_y = \mathrm{E}\left[(Y-\mu_y)(Y-\mu_y)^{\mathrm{T}}\right]$$

$$= A\mathrm{E}\left[(X-\mu_x)(X-\mu_x)^{\mathrm{T}}\right]A^{\mathrm{T}} \triangleq A\Sigma_x A^{\mathrm{T}}$$

1.1 概论

1.2 特征矢量和特征空间

1.3 随机矢量的描述

1.4 正态分布

1.4　　正态分布

正态分布的基本性、重要性

- 许多随机变量（近似）服从正态分布
- 是许多重要概率分布的极限分布
- 许多分布是正态分布的函数
- 有许多优良的性质
- 数学上便于运算

1.4.1 正态分布的定义

- 一维随机变量的正态分布

正态随机变量 x 的概率密度函数由期望 μ 和方差 σ^2 就可完全确定，定义为：

$$p(x)=\frac{1}{\sqrt{2\pi}\sigma}\exp\left[-\frac{(x-\mu)^2}{2\sigma^2}\right]$$

正态随机变量记为 $x \sim N(\mu,\sigma^2)$ 或 $p(x) \sim N(\mu,\sigma^2)$

1.4 正态分布

- 多变量正态分布

n 元正态分布的随机矢量 $\boldsymbol{X}=(X_1,X_2,\cdots,X_n)^{\mathrm{T}}$ 的概率密度函数定义为：

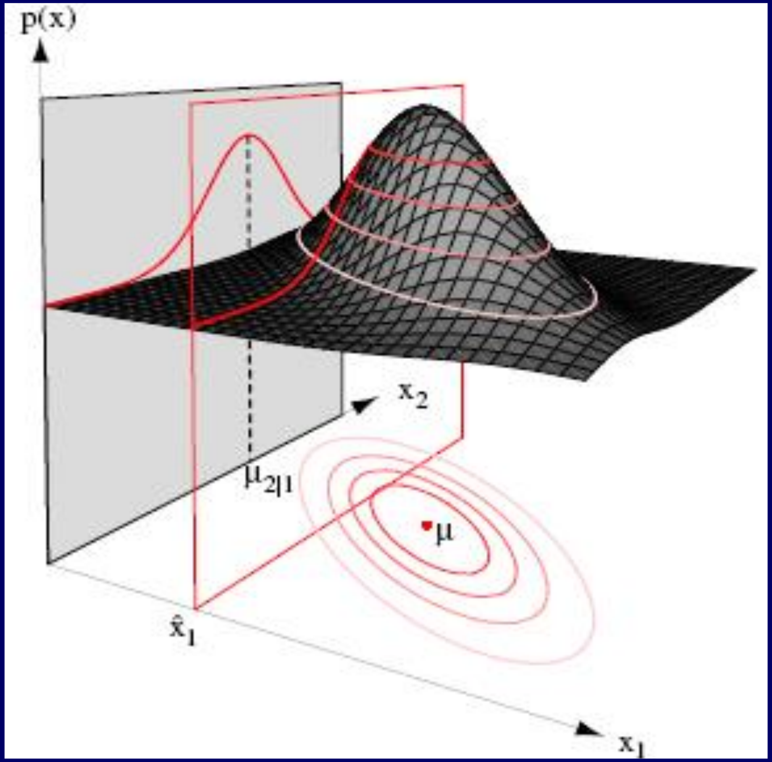
$$p(x_1,x_2,\cdots,x_n) \triangleq p(\boldsymbol{x})$$

$$=\frac{1}{(2\pi)^{n/2}\left|\boldsymbol{\Sigma}\right|^{1/2}}\exp\left[-\frac{1}{2}(\boldsymbol{x}-\boldsymbol{\mu})^{\mathrm{T}}\boldsymbol{\Sigma}^{-1}(\boldsymbol{x}-\boldsymbol{\mu})\right]$$

1.4.2 正态分布的性质

- 正态分布完全由μ和Σ确定，记 $X \sim N(\mu, \Sigma)$
- 等概率密度点的轨迹为一超椭球面
- 不相关与独立是等价的
- 多元正态随机矢量X 的边缘概率密度和条件概率密度仍是正态分布
- X 的线性变换仍为多元正态随机矢量
- X 的分量的线性组合是一多元正态随机变量

1.4　正态分布



二维正态分布的等概率密度点

轨迹、边缘概率密度函数示意图

作业

1. 试证明，对于正态分布，不相关与独立是等价的。
2. 试证明，多元正态随机矢量的线性变换仍为多元正态随机矢量。
3. 试证明，多元正态随机矢量X的分量的线性组合是一正态随机变量。

谢 谢！