第四章 随机变量的数字特征

§1 数学期望

§2 方差

§3 协方差及相关系数

§4 矩、协方差矩阵

§5 本章小结

方差的性质：

1. 设C是常数，则有$D(C) = 0$

证明：

$$D(C) = E\{[C - E(C)]^2\} = 0$$

2. 设X是一个随机变量，C是常数，则有$D(CX) = C^2 D(X)$

证明：

$$\mathrm{E}(CX) = E(C^2 x^2) - [E(Cx)]^2$$
$$= C^2\{E(x^2) - [E(x)]^2\} = C^2 D(X)$$

3. 设X,Y是两个随机变量，则有$D(X+Y)=D(X)+D(Y)+2E\{[X-E(X)][Y-E(Y)]\}$。特别，若$X,Y$相互独立，则有$D(X+Y)=D(X)+D(Y)$

证明：

$$
\begin{aligned}
D(X+Y) &= E\{[(X+Y)-E(X+Y)]^2\} \\
&= E\{[(X-E(X))+(Y-E(Y))]^2\} \\
&= E\{[X-E(X)]^2\}+E\{[Y-E(Y)]^2\} \\
&\quad +2E\{[X-E(X)][Y-E(Y)]\}
\end{aligned}
$$

若X,Y相互独立，$X-E(X)$与$Y-E(Y)$相互独立，有

$$
\begin{aligned}
&E\{[X-E(X)][Y-E(Y)]\} \\
&\quad = E[X-E(X)]E[Y-E(Y)]=0
\end{aligned}
$$

于是$D(X+Y)=D(X)+D(Y)$

将上述三项性质，若X，Y相互独立， a,b,c 是常数，则有：

$$D(aX+bY+c)=a^2D(X)+b^2D(Y)$$

这一性质可推广至任意有限个随机变量线性组合的情况

4. $D(X)=0 \Leftrightarrow P\big(X=E(X)\big)=1$

例6. 设活塞直径(cm) $X\sim N(22.40, 0.03^2)$，气缸直径 $Y\sim N(22.50, 0.04^2)$，X，Y相互独立，求任取一只活塞和一只气缸，求活塞能装入气缸的概率。

解：由正态分布和的分布关系，有

$$X-Y\sim N(-0.10, 0.05^2)$$

于是

$$P\{X\le Y\}=P\{X-Y\le 0\}$$
$$=P\left\{\frac{(X-Y)-(-0.10)}{0.05}\le\frac{0-(-0.10)}{0.05}\right\}$$
$$=\Phi(2)=0.9772$$

定理：令$X_1, X_2, \ldots, X_n$为iid随机变量，$\mu = E(X_i)$，$\sigma^2 = D(X_i)$定义样本均值为

$$\bar{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i$$

样本方差为

统计基础

$$S_n^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X}_n)^2$$

则

$$E(\bar{X}_n) = \mu, \qquad D(\bar{X}_n) = \frac{\sigma^2}{n}$$

$$E(S_n^2) = \sigma^2$$

兼听则明
偏信则暗

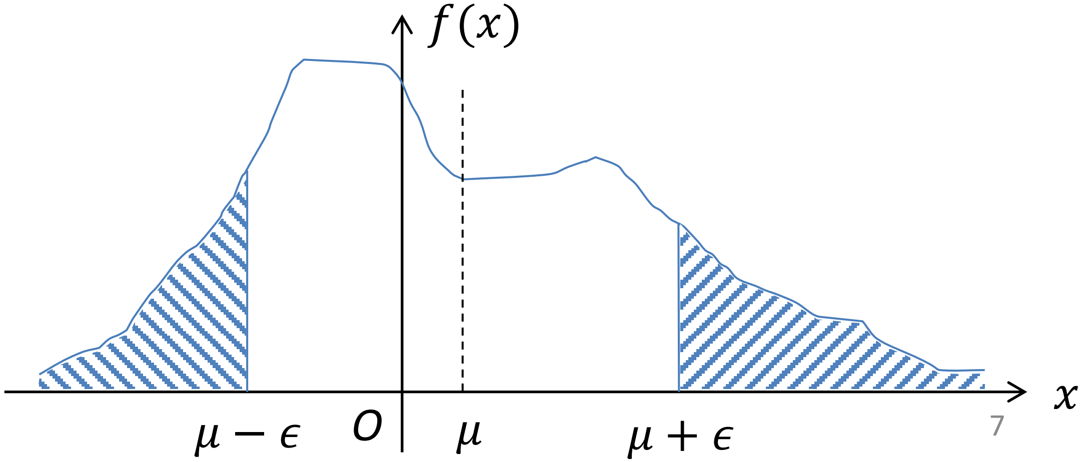
定理(切比雪夫(Chebyshev)不等式)

设随机变量X具有数学期望$E(X)=\mu$，方差$D(X)=\sigma^2$，则对于任意正数ϵ，不等式

$$P\{|X-\mu|\geq\epsilon\}\leq\frac{\sigma^2}{\epsilon^2}$$

$$\left(\text{等价地：}P\{|X-\mu|<\epsilon\}\geq 1-\frac{\sigma^2}{\epsilon^2}\right)$$

成立。



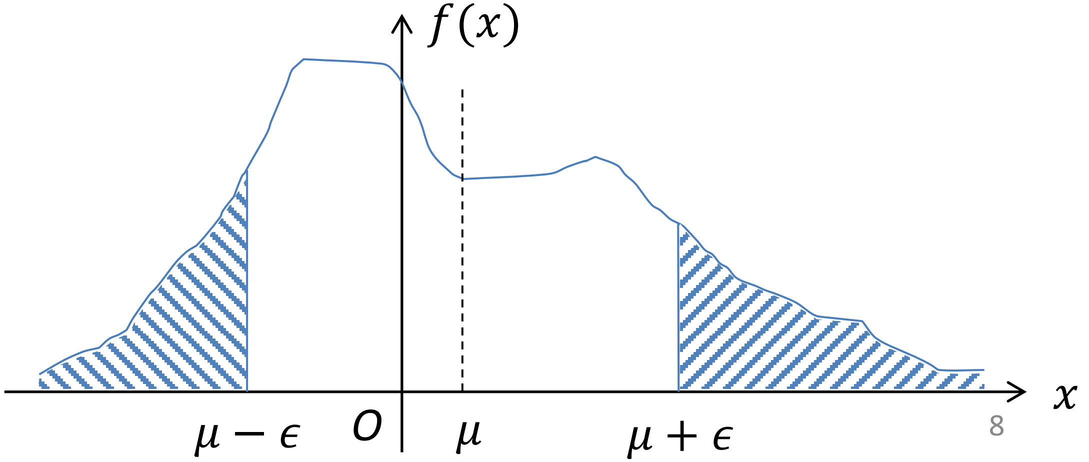
证明：

$$P\{|X-\mu| \geq \epsilon\}$$

$$= \int_{|x-\mu|\geq\epsilon} f(x)dx \leq \int_{|x-\mu|\geq\epsilon} \frac{|x-\mu|^2}{\epsilon^2} f(x)dx$$

$$\leq \frac{1}{\epsilon^2} \int_{-\infty}^{+\infty} (x-\mu)^2 f(x)dx$$

$$= \frac{\sigma^2}{\epsilon^2}$$



由切比雪夫(Chebyshev)不等式证下面的性质

$$D(X) = 0 \Leftrightarrow P(X = E(X)\ ) = 1$$

证明：由方差定义，若X=Const, 则显然，D(X)=0

若D(X)=0，假设$P(X = E(X)\ ) < 1$，即存在$\epsilon > 0$，使得$P(|X - E(X)| \geq \epsilon\ ) > 0$，但注意到切比雪夫不等式有

$$P(|X - E(X)| \geq \epsilon\ ) \leq \frac{D(X)}{\epsilon^2} = 0$$

与假设矛盾，于是$P(X = E(X)\ ) = 1$

§3 协方差(Covariance)及相关系数

对于二维随机变量(X,Y)，除了需要讨论X与Y的数学期望和方差外，还需讨论X与Y之间相互关系的数字特征

协方差的定义：

量$E\{[X-E(X)][Y-E(Y)]\}$称为随机变量X与Y的协方差，记为$\mathrm{Cov}(X,Y)$，即

$$\mathrm{Cov}(X,Y)= E\{[X-E(X)][Y-E(Y)]\}$$

称

$$\rho_{XY}=\frac{\mathrm{Cov}(X,Y)}{\sqrt{D(X)}\sqrt{D(Y)}}$$

为随机变量X与Y的相关系数，ρ_{XY}无量纲，有时又称为标准协方差

协方差的性质：

1. $\mathrm{Cov}(X,Y) = \mathrm{Cov}(Y,X)$
2. $\mathrm{Cov}(X,Y) = E(XY) - E(X)E(Y)$

证：$\mathrm{Cov}(X,Y) = E\{[X-E(X)][Y-E(Y)]\}$

$$= E\{XY - YE(X) - XE(Y) + E(X)E(Y)\}$$
$$= E(XY) - 2E(X)E(Y) + E(X)E(Y)$$
$$= E(XY) - E(X)E(Y)$$

3. $\mathrm{Cov}(aX,bY) = ab\mathrm{Cov}(X,Y)$ a,b是常数
4. $\mathrm{Cov}(X_1+X_2,Y) = \mathrm{Cov}(X_1,Y) + \mathrm{Cov}(X_2,Y)$

注意，上述公式中若X=Y，则退化为

1. $\mathrm{Cov}(X,X) = D(X)$
2. $\mathrm{Cov}(X,X) = E(X^2) - [E(X)]^2$
3. $\mathrm{Cov}(aX,bX) = abD(X)$ a,b是常数

协方差与随机变量和的方差

$$
\begin{aligned}
D(X+Y) &= E\{[(X+Y)-E(X+Y)]^2\} \\
&= E\left\{\left[\left(X-E(X)\right)+\left(Y-E(Y)\right)\right]^2\right\} \\
&= E\{[X-E(X)]^2\}+E\{[Y-E(Y)]^2\} \\
&\qquad +2E\{[X-E(X)][Y-E(Y)]\} \\
&= D(X)+D(Y)+2\mathrm{Cov}(X,Y)
\end{aligned}
$$

若X与Y相互独立，则有

$$
\mathrm{Cov}(X,Y)=E(XY)-E(X)E(Y)=0
$$

例1.设 $(X,Y)\sim N(\mu_1,\mu_2,\sigma_1^2,\sigma_2^2,\rho)$的正态分布，求X与Y的相关系数

解：

$$f(x,y)=\frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}}e^{-\frac{1}{2(1-\rho^2)}\left[\frac{(x-\mu_1)^2}{\sigma_1^2}-\frac{2\rho(x-\mu_1)(y-\mu_2)}{\sigma_1\sigma_2}+\frac{(y-\mu_2)^2}{\sigma_2^2}\right]}$$

由前面第三章第二节例10有

$$f_X(x)=\frac{1}{\sqrt{2\pi}\sigma_1}e^{-\frac{(x-\mu_1)^2}{2\sigma_1^2}}$$

$$f_Y(y)=\frac{1}{\sqrt{2\pi}\sigma_2}e^{-\frac{(y-\mu_2)^2}{2\sigma_2^2}}$$

由本章第一、第二节有

$$E(X)=\mu_1, E(Y)=\mu_2, D(X)=\sigma_1^2, D(Y)=\sigma_2^2$$

令$\mathrm{C}=\frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}}$，由协方差定义有

$$\mathrm{Cov}(X,Y)=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}(x-\mu_1)(y-\mu_2)f(x,y)dx\,dy$$

$$=C\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}(x-\mu_1)(y-\mu_2)$$

$$\cdot e^{-\frac{1}{2(1-\rho^2)}\left[\frac{(x-\mu_1)^2}{\sigma_1^2}-\frac{2\rho(x-\mu_1)(y-\mu_2)}{\sigma_1\sigma_2}+\frac{(y-\mu_2)^2}{\sigma_2^2}\right]}dxdy$$

$$=C\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}(x-\mu_1)(y-\mu_2)$$

$$\cdot e^{-\frac{(x-\mu_1)^2}{2\sigma_1^2}}e^{-\frac{1}{2(1-\rho^2)}\left[\frac{y-\mu_2}{\sigma_2}-\rho\frac{x-\mu_1}{\sigma_1}\right]^2}dy\,dx$$

做变量代换$t=\frac{1}{\sqrt{(1-\rho^2)}}\left[\frac{y-\mu_2}{\sigma_2}-\rho\frac{x-\mu_1}{\sigma_1}\right], u=\frac{x-\mu_1}{\sigma_1}$

$$\mathrm{Cov}(X,Y)=\frac{1}{2\pi}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\left(\sigma_1\sigma_2\sqrt{1-\rho^2}tu+\rho\sigma_1\sigma_2u^2\right)$$

$$\cdot e^{-\frac{t^2+u^2}{2}}dt\,du$$

$$=\frac{\rho\sigma_1\sigma_2}{2\pi}\left(\int_{-\infty}^{\infty}u^2e^{-\frac{u^2}{2}}du\right)\left(\int_{-\infty}^{\infty}e^{-\frac{t^2}{2}}dt\right)$$

$$+\frac{\sigma_1\sigma_2\sqrt{1-\rho^2}}{2\pi}\left(\int_{-\infty}^{\infty}ue^{-\frac{u^2}{2}}du\right)\left(\int_{-\infty}^{\infty}te^{-\frac{t^2}{2}}dt\right)$$

$$=\frac{\rho\sigma_1\sigma_2}{2\pi}\sqrt{2\pi}\sqrt{2\pi}=\rho\sigma_1\sigma_2$$

$$\rho_{XY}=\frac{Cov(X,Y)}{\sigma_1\sigma_2}=\rho$$

1. 二维正态分布密度函数中，参数ρ代表了X与Y的相关系数
2. 相关系数为零等价于X与Y相互独立

例2. 已知随机变量X, Y分别服从$N(1,3^2)$, $N(0,4^2)$，$\rho_{XY}=-1/2$，设$Z=X/3+Y/2$

求：(1) Z的数学期望和方差，(2) X与Z的相关系数

(3) X与Z是否独立？为何？

解：(1) $E(Z)=\frac{1}{3}E(X)+\frac{1}{2}E(Y)=\frac{1}{3}$

(2)

$$D(Z)=D(\frac{X}{3})+D(\frac{Y}{2})+2Cov(\frac{X}{3},\frac{Y}{2})$$
$$=\frac{1}{9}D(X)+\frac{1}{4}D(Y)+\frac{1}{3}Cov(X,Y)$$
$$=1+4+\frac{1}{3}\rho_{XY}\sqrt{D(X)}\sqrt{D(Y)}=3$$

(2)

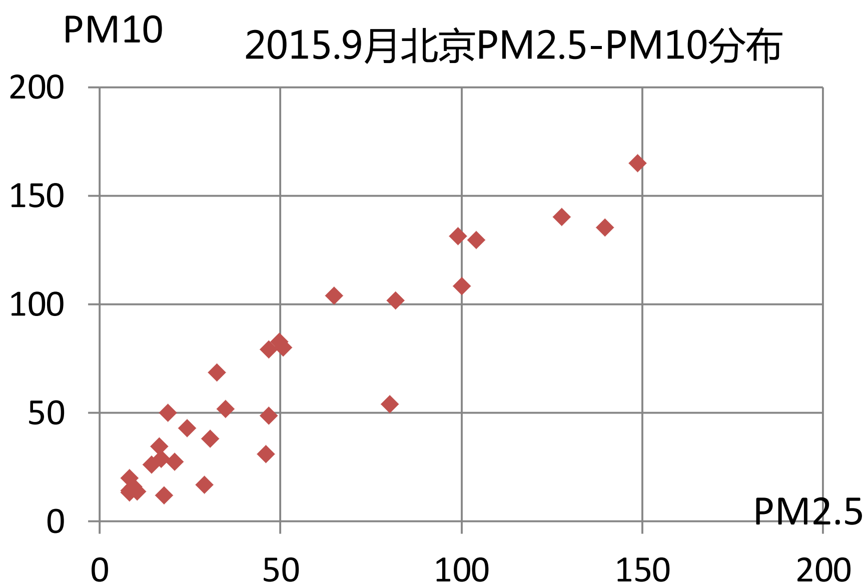
$$
\begin{aligned}
\mathrm{Cov}(X,Z) &= \mathrm{Cov}\left(X,\frac{X}{3}+\frac{Y}{2}\right)\\
&= \frac{1}{3}\mathrm{Cov}(X,X)+\frac{1}{2}\mathrm{Cov}(X,Y)\\
&= \frac{1}{3}D(X)+\frac{1}{2}\rho_{XY}\sqrt{D(X)}\sqrt{D(Y)}\\
&= 3-3=0
\end{aligned}
$$

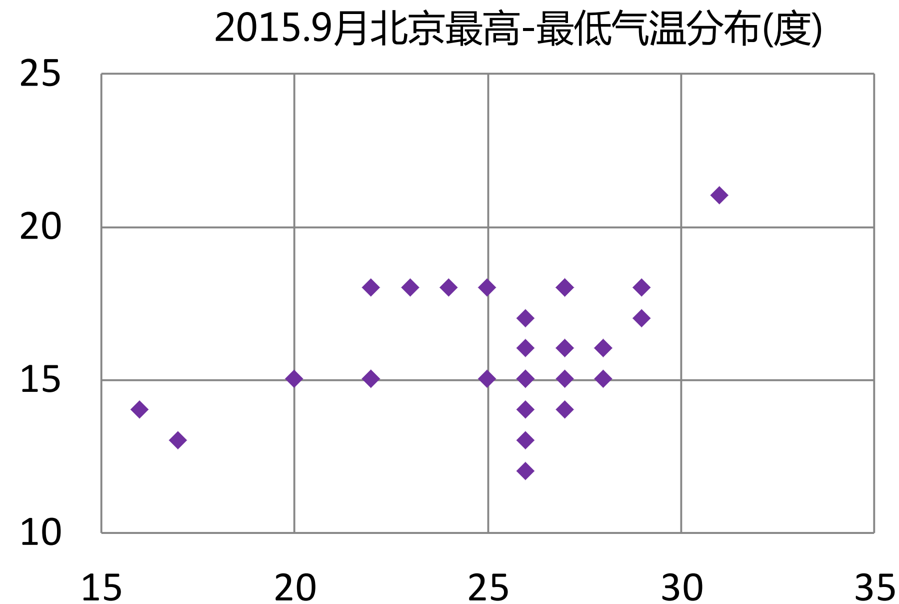
于是

$$\rho_{XY}=\frac{\mathrm{Cov}(X,Z)}{\sqrt{D(X)}\sqrt{D(Z)}}=0$$

(3) 注意到X和Z均为正态分布，于是X与Z是相互独立的

方差、协方差与相关系数



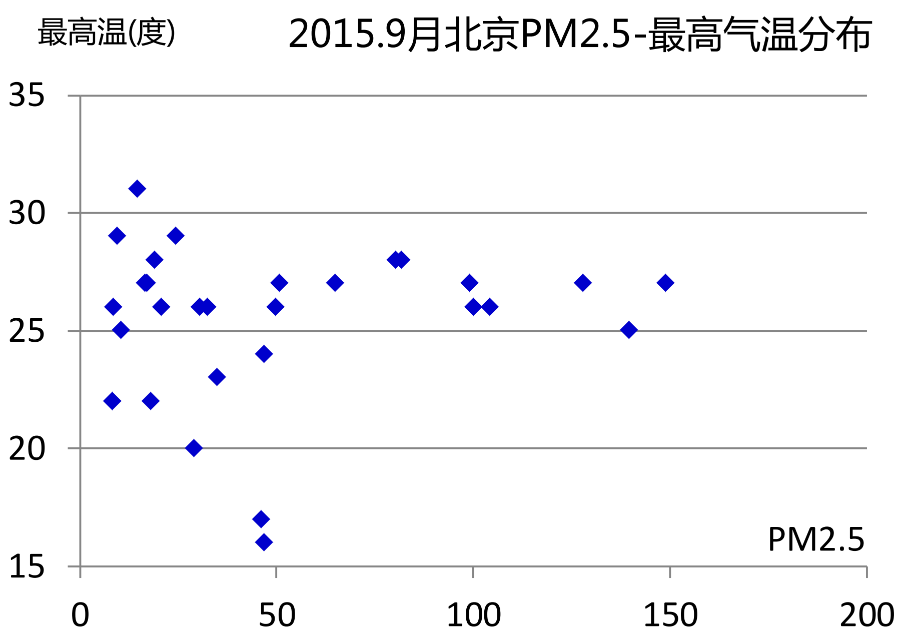


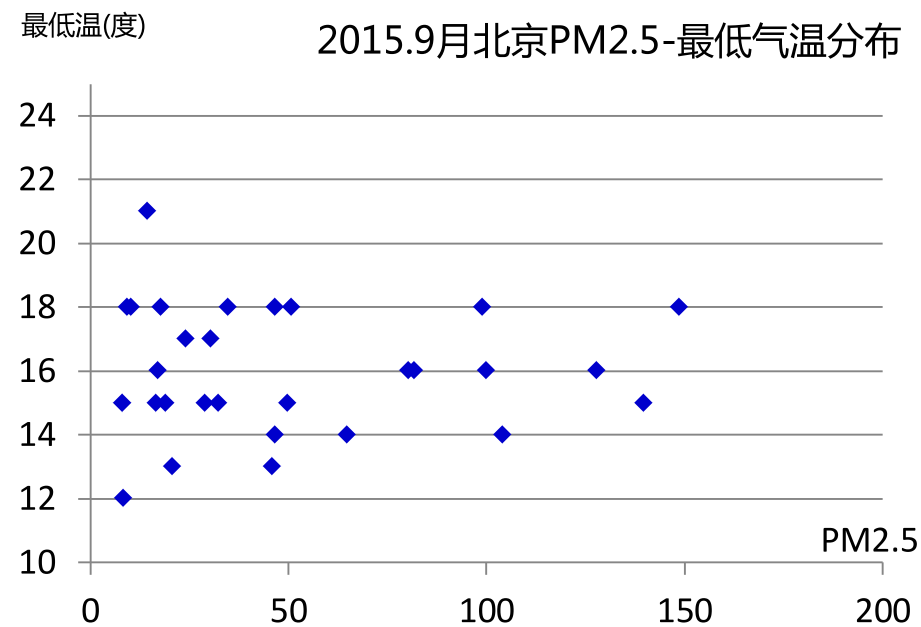
$$\mathrm{Var_{PM2.5}} = 1747.489$$
$$\mathrm{Var_{PM10}} = 2081.414$$
$$\mathrm{Cov} = 1788.025$$
$$\rho = 0.9375$$

$$\mathrm{Var_{Hi}} = 11.333$$
$$\mathrm{Var_{Lo}} = 3.895$$
$$\mathrm{Cov} = 2.632$$
$$\rho = 0.3962$$

PM2.5/PM数据源于：http://www.aqistudy.cn
气温数据源于：http://lishi.tianqi.com/

方差、协方差与相关系数



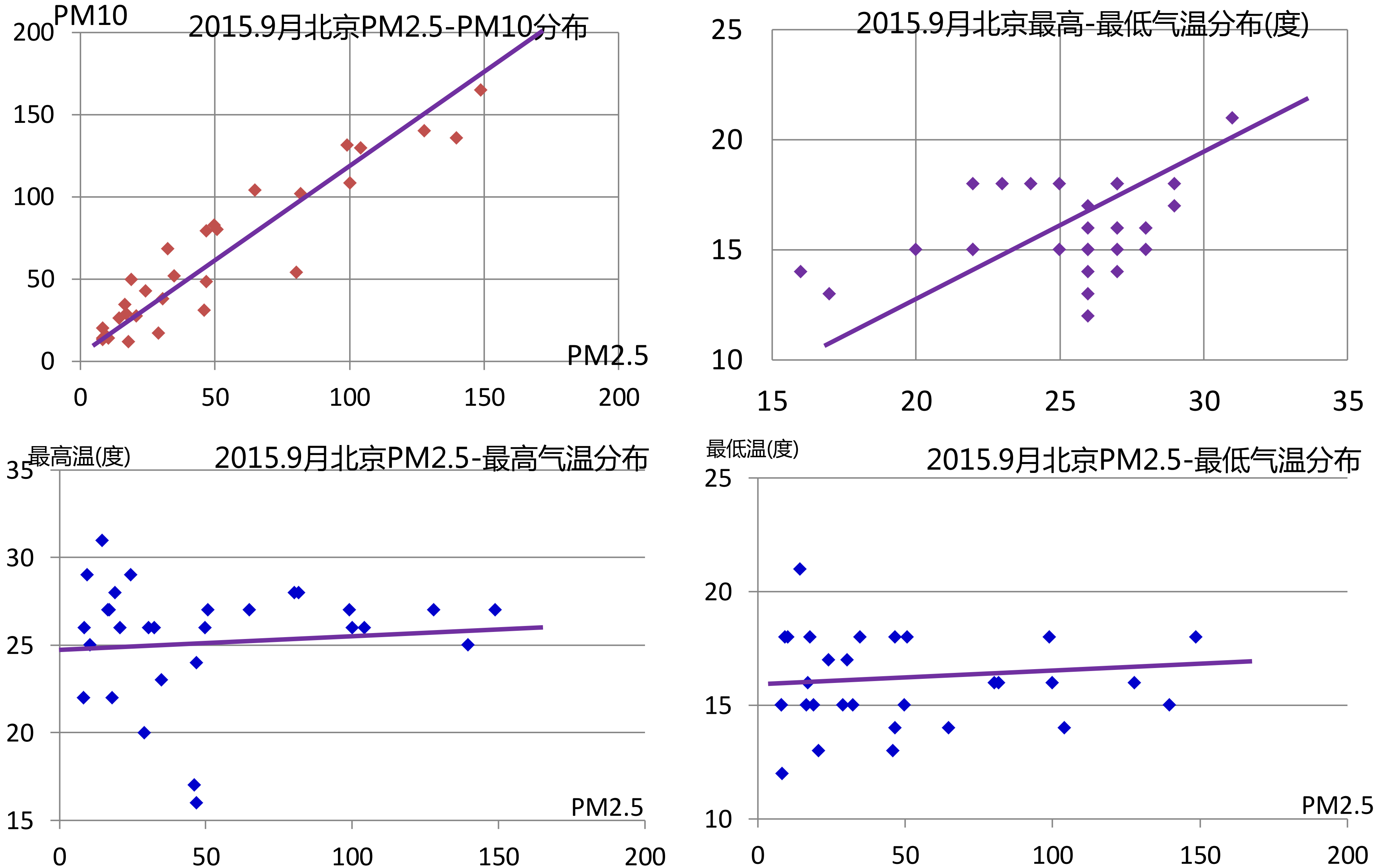


$$\mathrm{Var}_{\mathrm{PM2.5}} = 1747.489$$
$$\mathrm{Var}_{\mathrm{Hi}} = 11.333$$
$$\mathrm{Cov} = 17.647$$
$$\rho = 0.1254$$

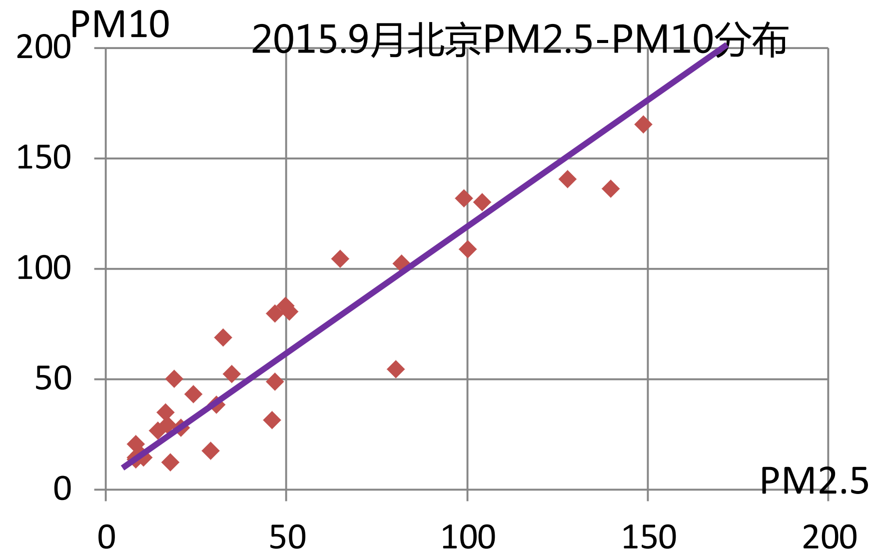
$$\mathrm{Var}_{\mathrm{PM2.5}} = 1747.489$$
$$\mathrm{Var}_{\mathrm{Lo}} = 3.895$$
$$\mathrm{Cov} = 1.550$$
$$\rho = 0.0188$$

PM2.5/PM数据源于：http://www.aqistudy.cn
气温数据源于：http://lishi.tianqi.com/

相关系数的几何意义


2015.9月北京PM2.5-PM10分布
PM10
200
150
100
50
0
0
50
100
150
200
PM2.5
2015.9月北京最高-最低气温分布(度)
25
20
15
10
15
20
25
30
35
2015.9月北京PM2.5-最高气温分布
最高温(度)
35
30
25
20
15
0
50
100
150
200
PM2.5
2015.9月北京PM2.5-最低气温分布
最低温(度)
25
20
15
10
0
50
100
150
200
PM2.5

相关系数的几何意义



随机变量X与Y是否可以通过线性拟合(回归)加以逼近

$$\min_{a,b}\left\{e = E\left[\left(Y-(a+bX)\right)^2\right]\right\} \qquad (*1)$$

求关于a,b的偏导，并令它们为零，有

$$\begin{cases} \dfrac{\partial e}{\partial a} = 2a + 2bE(X) - 2E(Y) = 0 \\ \dfrac{\partial e}{\partial b} = 2bE(X^2) - 2E(XY) + 2aE(X) = 0 \end{cases}$$

解得： $a=E(Y)-\frac{E(X)\mathrm{Cov}(X,Y)}{D(X)}, b=\frac{\mathrm{Cov}(X,Y)}{D(X)}$

代回(*1)

$$e=E\left[\left(Y-E(Y)-\frac{\mathrm{Cov}(X,Y)}{D(X)}\big(X-E(X)\big)\right)^2\right]$$

$$=D(Y)-\frac{[\mathrm{Cov}(X,Y)]^2}{D(X)}=D(Y)(1-\rho_{XY}^2)$$

几何解释：

1. 当$|\rho_{XY}|$较大时，e较小，X,Y的线性关系较为紧密$\rho_{XY}=1$，$e=0$，X与Y以概率1线性相关
2. 当$|\rho_{XY}|$较小时，e较大，X,Y的线性关系较差，$\rho_{XY}=0$时，X与Y完全不相关

随机变量的不相关与独立

定义：若$\rho_{XY} = 0$，则称随机变量X与Y不相关

注意：是否相关只是针对线性关系而言的，X与Y相互独立则是相对一般关系而言的

随机变量X与Y不相关即$\rho_{XY} = 0$的等价条件有：

(1) 由$\rho_{XY} = 0 \ \Longrightarrow \mathrm{Cov}(X,Y) = 0$

(2) $\Longrightarrow\ E(XY) - E(X)E(Y) = 0$，即$E(XY) = E(X)E(Y)$

(3) $\Longrightarrow D(X+Y) = D(X) + D(Y) + 2E\{[X - E(X)][Y - E(Y)]\} = D(X) + D(Y)$

X与Y相互独立 $\Longrightarrow$ X与Y一定不相关，反之则不然

例3. 设θ服从$[0,2\pi]$上的均匀分布，$\xi = \cos\theta$，$\eta = \cos(\theta + a)$，这里a是常数，求ξ和η的相关系数

解：

$$E(\xi) = \frac{1}{2\pi}\int_0^{2\pi} \cos x \, dx = 0$$

$$E(\eta) = \frac{1}{2\pi}\int_0^{2\pi} \cos(x + a)\, dx = 0$$

$$D(\xi) = \frac{1}{2\pi}\int_0^{2\pi} \cos^2 x \, dx = \frac{1}{2\pi}\left(\frac{x}{2} + \frac{1}{2}\sin x \cos x\right)\Bigg|_0^{2\pi} = \frac{1}{2}$$

$$D(\eta) = \frac{1}{2\pi}\int_0^{2\pi} \cos^2(x + a)\, dx$$

$$= \frac{1}{2\pi}\left(\frac{x}{2} + \frac{1}{2}\sin x \cos x\right)\Bigg|_a^{2\pi+a} = \frac{1}{2}$$

$$\operatorname{cov}(\xi\eta)=\frac{1}{2\pi}\int_0^{2\pi}\cos x\cos(\,x+a)\,dx=\frac{1}{2}\cos a$$

于是有相关系数：

$$\rho=\frac{\operatorname{cov}(\xi\eta)}{\sqrt{D(\eta)}\sqrt{D(\eta)}}=\cos a$$

1. $a=0$时，$\rho=1$， $\xi=\eta$，存在线性关系
2. $a=\pi$时，$\rho=-1$， $\xi=-\eta$，存在线性关系
3. $a=\frac{\pi}{2}$或$a=\frac{3\pi}{2}$时，$\rho=0$， ξ与η不相关
   但$\xi^2+\eta^2=1$，因此ξ与η不独立

§4 矩、协方差矩阵

定义：设X和Y是随机变量，若$E\left(X^k\right)\ (k=1,2,\dots)$存在，则称它为X的k阶原点矩，简称k阶矩。

若$E\{[X-E(X)]^k\}\ (k=2,3,\dots)$存在，称它为X的k阶中心矩

若$E\left(X^kY^l\right)\ (k,l=1,2,\dots)$存在，称它为X和Y的$k+l$阶混合矩

若$E\{[X-E(X)]^k[Y-E(Y)]^l\}\ (k,l=1,2,\dots)$存在，称它为X和Y的$k+l$阶中心矩

数学期望E(X)是X的一阶原点矩

方差D(X)是X的二阶中心矩

协方差Cov(X,Y)是X与Y的二阶混合中心矩

三阶中心矩可以用来衡量随机变量的分布是否有偏

更高阶中心矩可以用来衡量随机变量的分布在均值附近的陡峭程度如何

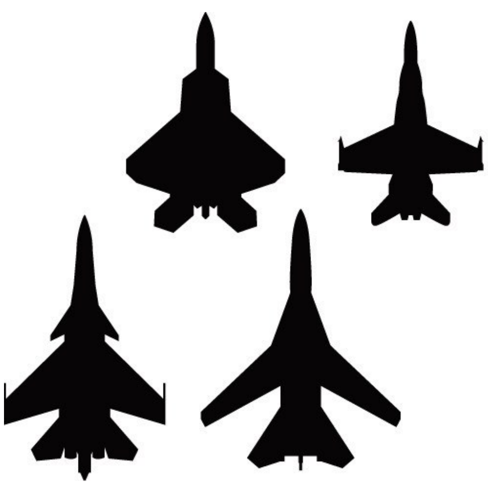
高阶矩很少使用

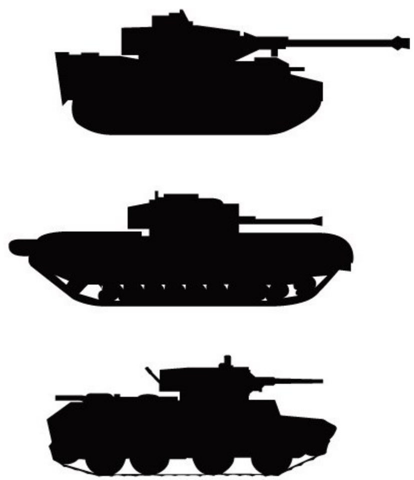
用于图像识别的Hu矩*

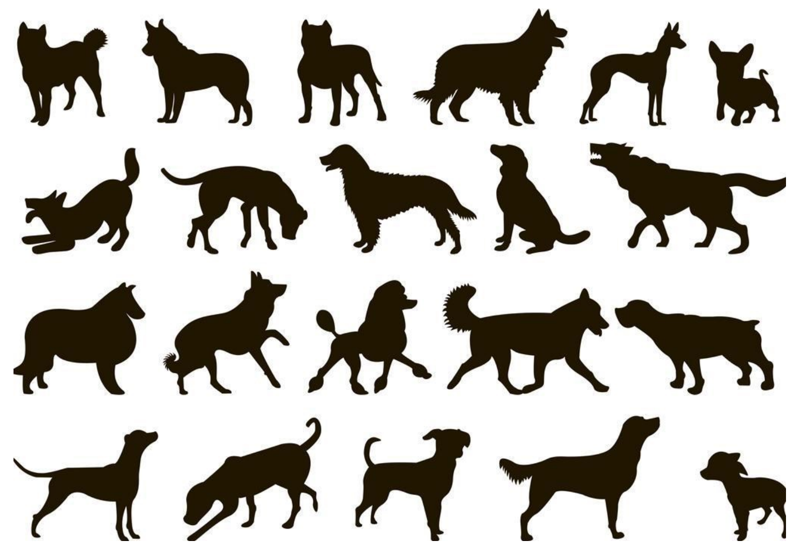
$$m_{pq}=\sum_{y=1}^{N}\sum_{x=1}^{M}x^{p}y^{q}f(x,y)\ ,$$

$$\mu_{pq}=\sum_{y=1}^{N}\sum_{x=1}^{M}(x-\bar{x})^{p}(y-\bar{y})^{q}f(x,y)\ ,p,q=0,1,\ldots$$

$$\eta_{pq}=\frac{\mu_{pq}}{\mu_{00}^{\rho}},\rho=\frac{p+q}{2}+1$$







* MK Hu , Visual pattern recognition by moment invariants, IRE Transactions on Information Theory, 8(2), 1962

由二阶和三阶归一化中心矩构造了7个不变矩

$$M_1 = \eta_{20} + \eta_{02}$$

$$M_2 = (\eta_{20} - \eta_{02})^2 + 4\eta_{11}$$

$$M_3 = (\eta_{30} - 3\eta_{12})^2 + (3\eta_{21} - \eta_{03})^2$$

$$M_4 = (\eta_{30} + \eta_{12})^2 + (\eta_{21} + \eta_{03})^2$$

$$\begin{aligned} M_5 = {} & (\eta_{30} - 3\eta_{12})(\eta_{30} + \eta_{12})[(\eta_{30} + \eta_{12})^2 - 3(\eta_{21} + \eta_{03})^2] \\ & + (3\eta_{21} - \eta_{03})(\eta_{21} + \eta_{03})[3(\eta_{30} + \eta_{12})^2 \\ & - (\eta_{21} + \eta_{03})^2] \end{aligned}$$

$$\begin{aligned} M_6 = {} & (\eta_{20} - \eta_{02})\,[(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2] \\ & +4\eta_{11}(\eta_{30} + \eta_{12})(\eta_{21} + \eta_{03}) \end{aligned}$$

$$\begin{aligned} M_7 = {} & (3\eta_{21} - \eta_{03})(\eta_{30} + \eta_{12})[(\eta_{30} + \eta_{12})^2 - 3(\eta_{21} + \eta_{03})^2] \\ & - (\eta_{30} - 3\eta_{12})(\eta_{21} + \eta_{03})[3(\eta_{30} + \eta_{12})^2 \\ & - (\eta_{21} + \eta_{03})^2] \end{aligned}$$

协方差矩阵

设n维随机变量$(X_1, X_2, \dots, X_n)$的二阶混合中心矩，
$c_{ij} = Cov(X_i, X_j) = E\{[X_i - E(X_i)][X_j - E(X_j)]\}$,
$(i, j = 1,2, \dots, n)$都存在，则称矩阵

$$C = \begin{Bmatrix} c_{11} & c_{12} & \cdots & c_{1n} \\ c_{21} & c_{22} & \cdots & c_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ c_{n1} & c_{n2} & \cdots & c_{nn} \end{Bmatrix}$$

为n维随机变量的协方差矩阵

1. 协方差矩阵是一个对称非负定矩阵
2. 协方差矩阵可用来表示多维随机变量的概率密度，从而可通过协方差矩阵达到对多维随机变量的研究

设 $(X_1,X_2)\sim N(\mu_1,\mu_2,\sigma_1^2,\sigma_2^2,\rho)$的正态分布，即

$$f(x_1,x_2) = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} e^{-\frac{1}{2(1-\rho^2)}\left[\frac{(x_1-\mu_1)^2}{\sigma_1^2}-\frac{2\rho(x_1-\mu_1)(x_2-\mu_2)}{\sigma_1\sigma_2}+\frac{(x_2-\mu_2)^2}{\sigma_2^2}\right]}$$

引入 $\boldsymbol{X}=\begin{pmatrix}x_1\\x_2\end{pmatrix},\boldsymbol{\mu}=\begin{pmatrix}\mu_1\\\mu_2\end{pmatrix}$

协方差矩阵为：

$$\boldsymbol{C}=\begin{pmatrix}c_{11} & c_{12}\\ c_{21} & c_{22}\end{pmatrix}=\begin{pmatrix}\sigma_1^2 & \rho\sigma_1\sigma_2\\ \rho\sigma_1\sigma_2 & \sigma_2^2\end{pmatrix}$$

$$\boldsymbol{C}^{-\mathbf{1}}=\frac{1}{\det\boldsymbol{C}}\begin{pmatrix}\sigma_2^2 & -\rho\sigma_1\sigma_2\\ -\rho\sigma_1\sigma_2 & \sigma_1^2\end{pmatrix}$$

$$=\frac{1}{\sigma_1^2\sigma_2^2(\mathbf{1}-\boldsymbol{\rho}^{\mathbf{2}})}\begin{pmatrix}\sigma_2^2 & -\rho\sigma_1\sigma_2\\ -\rho\sigma_1\sigma_2 & \sigma_1^2\end{pmatrix}$$

注意到

$$(\boldsymbol{X}-\boldsymbol{\mu})^{T}\boldsymbol{C}^{-\mathbf{1}}(\boldsymbol{X}-\boldsymbol{\mu}) = \frac{1}{1-\rho^2}\left[\frac{(x_1-\mu_1)^2}{\sigma_1^2} - 2\rho\frac{(x_1-\mu_1)(x_2-\mu_2)}{\sigma_1\sigma_2} + \frac{(x_2-\mu_2)^2}{\sigma_2^2}\right]$$

$$f(x_1,x_2) = \frac{1}{2\pi(det\boldsymbol{C})^{\frac{1}{2}}}\exp\left[-\frac{1}{2}(\boldsymbol{X}-\boldsymbol{\mu})^{T}\boldsymbol{C}^{-\mathbf{1}}(\boldsymbol{X}-\boldsymbol{\mu})\right]$$

类似地，可以推广到n维

$$\boldsymbol{X}=\begin{pmatrix}x_1\\x_2\\\vdots\\x_n\end{pmatrix},\boldsymbol{\mu}=\begin{pmatrix}\mu_1\\\mu_2\\\vdots\\\mu_n\end{pmatrix},\boldsymbol{C}\text{是协方差矩阵}$$

$(X_1,X_2,\dots,X_n)$的概率密度定义为：

$$f(x_1,x_2,\dots,x_n)=\frac{1}{(2\pi)^{\frac{n}{2}}|\boldsymbol{C}|^{\frac{1}{2}}}\exp\left[-\frac{1}{2}(\boldsymbol{X}-\boldsymbol{\mu})^{\boldsymbol{T}}\boldsymbol{C}^{-\mathbf{1}}(\boldsymbol{X}-\boldsymbol{\mu})\right]$$

n维正态分布的性质

1. n维正态变量$(X_1, X_2, \dots, X_n)$的每一个分量X_i (i=1,2,...)都是正态变量；反之，若$X_1, X_2, \dots, X_n$都是正态变量，且相互独立，则$(X_1, X_2, \dots, X_n)$是n维正态变量；
2. n维正态变量$(X_1, X_2, \dots, X_n)$服从n维正态分布$\Leftrightarrow$ $X_1, X_2, \dots, X_n$的任意线性组合服从正态分布
3. 若$(X_1, X_2, \dots, X_n)$服从n维正态分布，设$Y_1, Y_2, \dots, Y_k$是X_j的线性函数，则$(Y_1, Y_2, \dots, Y_k)$也服从多维正态分布
4. 设$(X_1, X_2, \dots, X_n)$服从n维正态分布，则$X_1, X_2, \dots, X_n$相互独立$\Leftrightarrow$ $X_1, X_2, \dots, X_n$两两不相关

例：利用协方差进行识别

- 一种局部描述子，在纹理分类、物体检测等方面获得很好的应用
- 用d维特征(比如R,G,B和前2阶差分)的Covariance描述了感兴趣的区域
- 协方差矩阵不在欧式空间上，因此采用广义特征值表示的距离度量

Tuzel, Porikli, and Peter Meer， Region Covariance: A Fast Descriptor for Detection and Classification， ECCV 2006

作为区域描述子

给定区域R，其尺寸为W×H

$\{z_k\}_{k=1,2,\ldots,n} \in R$为R中的n个d维特征，采用dxd协方差矩阵作为特征描述

$$C_R = \frac{1}{n-1}\sum_{k=1}^{n}(z_k-\mu)(z_k-\mu)^T$$

μ是特征点的均值

特点

- 判别能力强
- 自然的特征融合方式
  - 对角线对应每个特征的方差
  - 非对角线对应特征间的相关性
  - 噪声可以用平均的方法滤除
- 特征维度低
  - 直方图特征维度：$\prod_{i=1}^{d} h_i$，其中h_i为特征i的Bin的数量
  - 原始数据的维数：$n \times d$
  - 协方差特征的维数：$(d^2 + d)/2$
- 一定尺度上的旋转和尺度不变

距离计算

协方差矩阵的距离不能有欧式空间计算

两个协方差矩阵的不相似度定义为[*]，

$$\rho(C_1, C_2) = \sqrt{\sum_{i=1}^{n} \ln^2 \lambda_i(C_1, C_2)}$$

其中$\lambda_i(C_1, C_2)$是C_1, C_2的广义特征值，由下式计算获得

$\lambda_i C_1 \mathbf{x}_i - C_2 \mathbf{x}_i = 0$，$\mathbf{x}_i$为第i个泛化的特征向量

$\rho(C_1, C_2)$满足距离的非负性、交换性和三角不等式

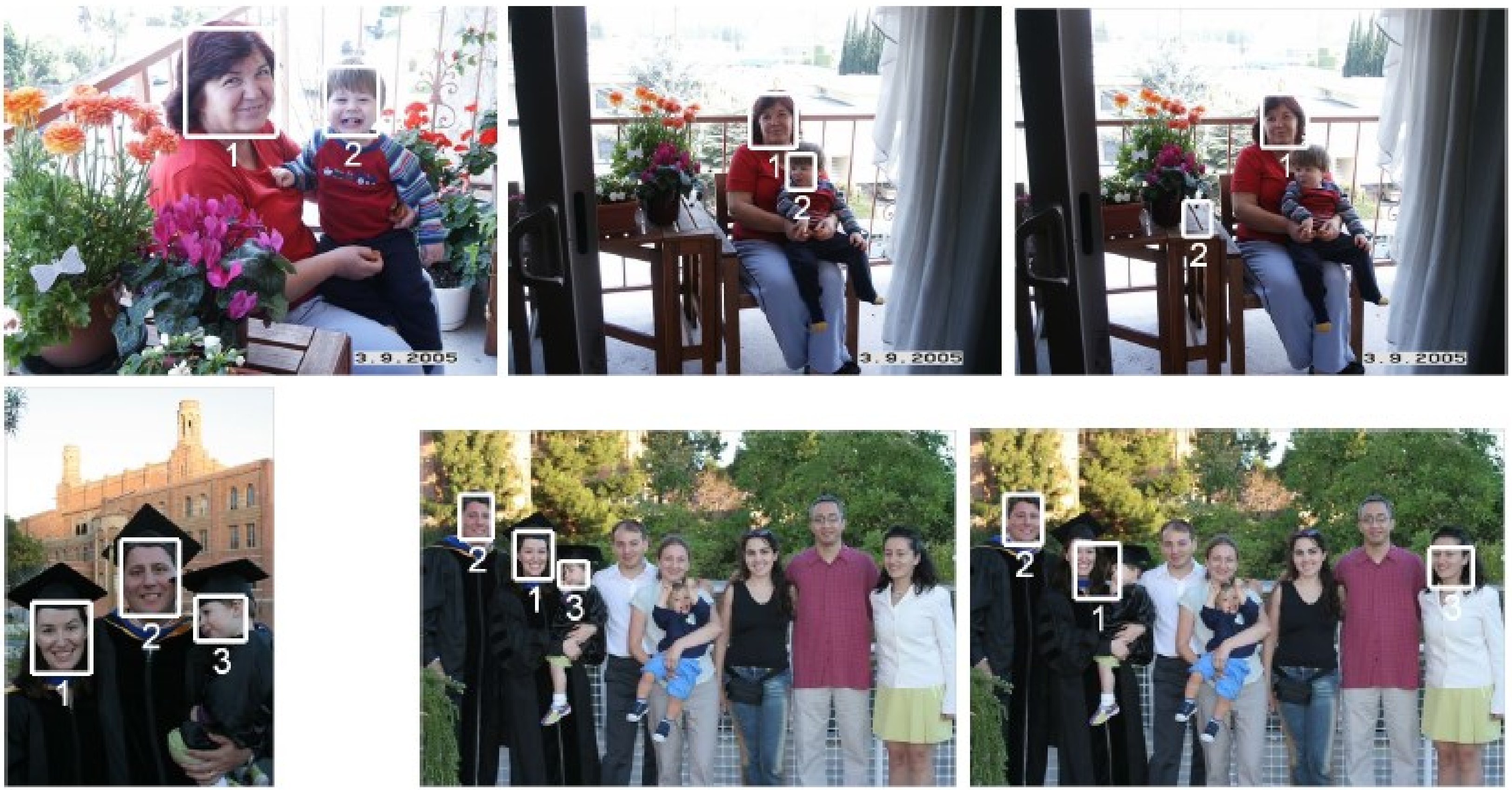
$\rho(C_1, C_2) \geq 0$，且$\rho(C_1, C_2) = 0$，当且仅当$C_1 = C_2$

$$\rho(C_1, C_2) = \rho(C_2, C_1)$$

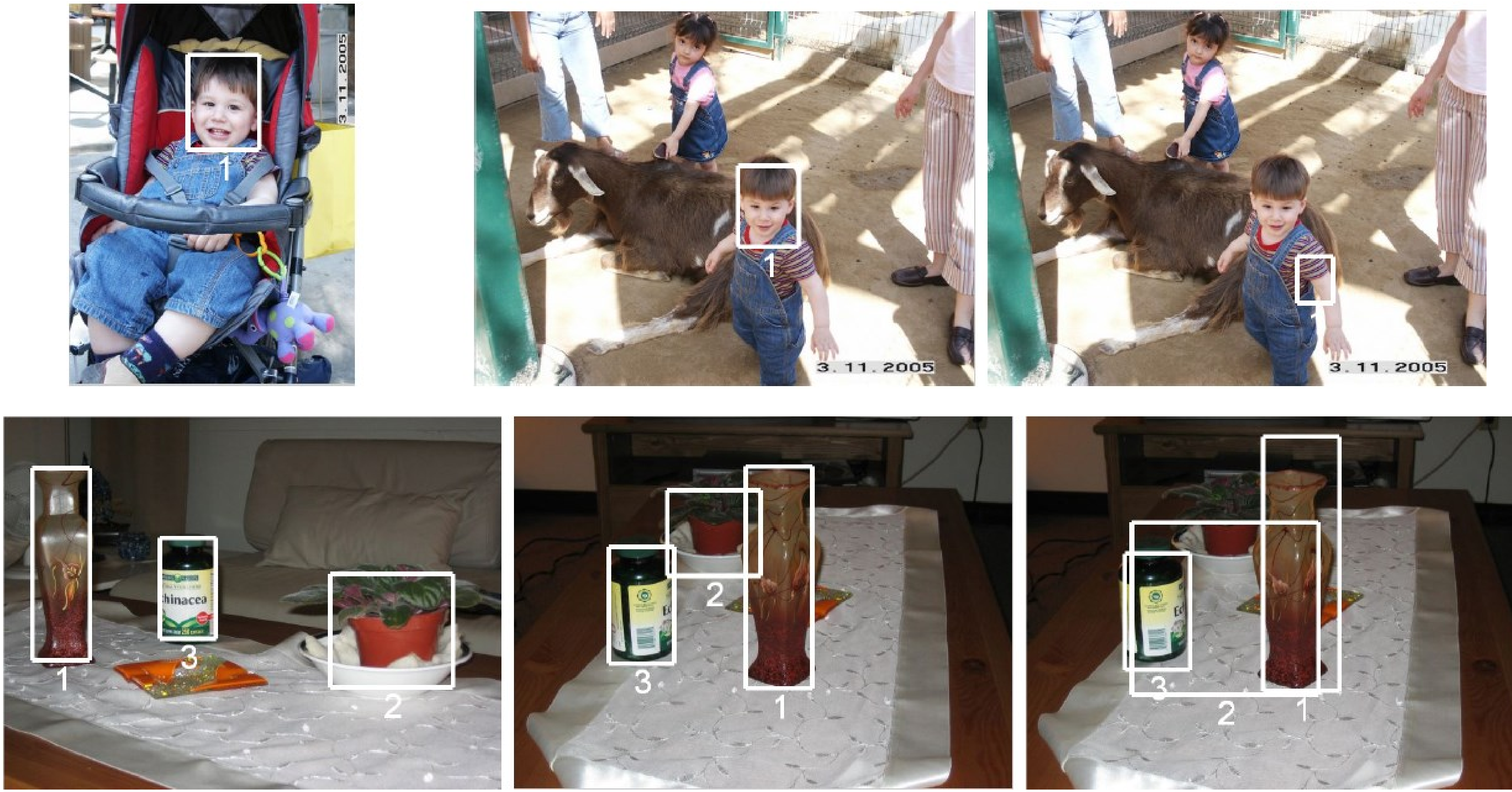
$$\rho(C_1, C_2) + \rho(C_1, C_3) > \rho(C_2, C_3)$$

[*] W. Forstner, B. Moonen. A metric for covariance matrices. Technical report, Dept. of Geodesy and Geoinformatics, Stuttgart University (1999)

目标检测示例



目标检测示例



目标检测示例

§5 本章小结

- 总体特性是关心概率的目的
- 从矩的角度理解多种数字特征
- 矩包含了多种数字特征
- 协方差矩阵包含了丰富的信息
- 契比雪夫不等式
- 不相关性与独立性之间的关系
- 正态随机变量的特征

作业

概率论与数理统计

pp. 117-118, #34，#35，#37