Multivariable Correlation

C_n^2个相关系数

$X=(X_1,X_2,\cdots,X_n)^T$ Joint Distribution ? $\Rightarrow$ Correlation Matrix

$R_X=E(XX^T)$ $\quad R_X(i,j)=E(X_i,X_j)$ $\quad R_X=R_X^T$

① Decorrelation (Whitening) 去相关性/白化 $\quad A\in\mathbb{R}^{n\times n}$, $Y=AX\in\mathbb{R}^n$

线性变换,去对角外相关

$E(Y_i,Y_j)=0\ (i\neq j)$

$\Rightarrow R_Y=E(AX)(AX)^T=AE(XX^T)A^T=AR_XA^T$

矩阵特征分解

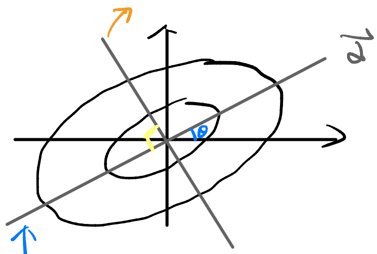
$R_X=\sum_{k=1}^{n}\lambda_k U_kU_k^T=U\Lambda U^T$, $U=(U_1,\cdots,U_k)$, $U\cdot U^T=U^T\cdot U=I$

$\Lambda=diag(\lambda_1,\cdots,\lambda_n)$, $\lambda_k\geq 0,\ \forall k$

$\Rightarrow U^TR_XU=\Lambda \quad\Rightarrow \boxed{A=U^T}$ ★

② Principal Component Analysis (PCA)

正交方向损失最小



在这根轴上分布比较开,
方差比较大

(Dimensional Deduction)
Compression

$\alpha\in\mathbb{R}^n$, $E\|Proj_\alpha X\|^2=E\left\|\frac{\alpha^TX}{\alpha^T\alpha}\cdot\alpha\right\|^2$

X投影到α方向上有最大方差

$=\frac{E|\alpha^TX|^2}{\|\alpha\|^4}\cdot\|\alpha\|^2=E\left(\frac{\alpha^T}{\|\alpha\|}\cdot X\right)^2$

找α

$\max_\alpha E\|Proj_\alpha X\|^2=\max_\alpha E\left(\frac{\alpha^T}{\|\alpha\|}X\right)^2$

$\Longleftrightarrow \max_\alpha E(\alpha^TX)^2$ $\quad$ s.t. $\|\alpha\|=1$ $\quad$ 约定投影到α上的α模为1

$E(\alpha^TX)^2=E(\alpha^TX)(X^T\alpha)=\alpha^TE(XX^T)\alpha=\alpha^TR_X\alpha$

$$\boxed{\max_\alpha \alpha^TR_X\alpha\ ,\ s.t.\ \|\alpha\|=1}$$

$L(\alpha,\lambda)=\alpha^TR_X\alpha+\lambda(\overset{\alpha^T\alpha}{\|\alpha\|^2}-1)$

求导

$\nabla_\alpha L(\alpha,\lambda)=2\cdot R_X\cdot\alpha-2\lambda\alpha=0 \iff R_X\alpha=\lambda\alpha$

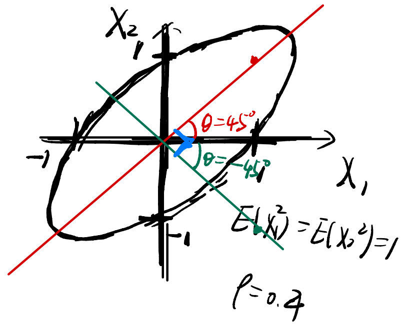
$\max_\alpha \alpha^TR_X\alpha=\lambda\alpha^T\alpha=\lambda \Rightarrow$ Eigenvector w.r.t largest Eigenvalue

特征值最大

对称矩阵特征向量性质:不同特征值对应的特征向量正交

以特征向量偏移角度 θ 分析相关性

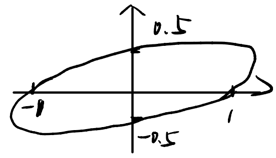
$E(X_1)=E(X_2)=0$, $\boxed{E(X_1^2)=E(X_2^2)=1}$, $E(X_1X_2)=\rho$, $X=(X_1,X_2)^T$



$R_X=\begin{pmatrix}1&\rho\\ \rho&1\end{pmatrix} \Rightarrow \lambda_1=1+\rho$, $\lambda_2=1-\rho \Rightarrow U_1=\begin{pmatrix}1\\1\end{pmatrix}$ $45°$ $\quad U_2=\begin{pmatrix}1\\-1\end{pmatrix}$ $-45°$ 角度固定

决定角度的是方差 不是相关系数 ρ

ρ 大了相关性强，X_1 与 X_2 结合"更近"，椭圆"更瘦"



$E(X_1^2)=1$
$E(X_2^2)=0.5$
$\rho=0.6$

③ Expansion $\quad Y=AX \Rightarrow X=A^{-1}Y \Rightarrow X=U\cdot Y=\sum_{k=1}^{n} U_k\cdot y_k$

$U_i^TU_j=0 \ (i\neq j)$ $\qquad E(Y_iY_j)=0 \ (i\neq j)$ $\qquad$ (Biorthogonal Expansion)

orthogonal $\qquad$ Orthogonal

系数有正交性 $\qquad$ 基也有正交性

Karhunan-Loeve Expansion (K-L 展开)

$\exp(i\omega_k t)$ → 一定区域正交，但系数正交性无法保证

$X(t)=\sum_{k=-\infty}^{\infty}\alpha_k\phi_k(t)$

展开，将随机性和过程性解耦
寻找比傅里叶展开(单正交)更高级的
双正交展开(基正交+系数也正交)

Assume $\{\phi_k\}$ orthogonal

假定 $\{\phi_k\}$ 在区间 I 上归一化 $\int_I\phi_k^2(t)dt=1$ $\qquad \int_I\phi_i(t)\phi_j(t)dt=0$

$\alpha_k=\int_I X(t)\phi_k(t)dt$

$\Rightarrow \boxed{E(\alpha_i\alpha_j)}=E\left(\int_I X(t)\phi_i(t)dt\int_I X(s)\phi_j(s)ds\right)$

$=\int_I\int_I E(X(t)X(s))\phi_i(t)\phi_j(s)dtds=\int_I\int_I R_X(t,s)\phi_i(t)\phi_j(s)dtds=0 \ (i\neq j)$

$\sum\sum \quad R_X(m,n) \quad \phi_i(m) \quad \phi_j(n)$

1. 积分看作求和 2. 二元函数看作矩阵 3. 一元函数看作矢量

$\Rightarrow \sum_m\sum_n R_X(m,n)\phi_i(m)\phi_j(n)$

$\phi_i^TR_X\phi_j=\phi_i^T(\lambda)\phi_j)=0$ → 连续版本下特征函数

对称

$\int_I R_X(t,s)\phi_i(s)ds=\lambda_i\phi_i(t)$ $\quad \forall i$ (Integral Equation) → 积分方程

$\sum_n R_X(m,n)\phi_i(n)=\lambda_i\phi_i(m) \Leftrightarrow R_X\cdot\phi_i=\lambda_i\phi_i$

→ 对称矩阵特征向量正交(上证)

1. Karhunen-Loève 展开

K-L 展开用于将一个随机过程 $X(t)$ 以一组正交基函数 $\{\phi_k(t)\}$ 进行展开：

$$X(t) = \sum_{k=1}^{\infty} \alpha_k \phi_k(t)$$

其中 $\{\phi_k(t)\}$ 是一组正交基函数，而 α_k 是对应的随机系数。

2. 计算系数 α_k

为了计算 α_k，我们利用**正交投影**的思想。

1. **两边同时与 $\phi_k(t)$ 相乘并在定义域 I 上积分**：

$$\int_I X(t)\phi_k(t)dt = \int_I \left(\sum_{j=1}^{\infty} \alpha_j \phi_j(t) \right) \phi_k(t)dt$$

2. **交换求和与积分**（在合适的数学条件下是成立的）：

$$\sum_{j=1}^{\infty} \alpha_j \int_I \phi_j(t)\phi_k(t)dt$$

3. **利用正交性**：由于 $\{\phi_k(t)\}$ 是一组正交基，满足：

$$\int_I \phi_j(t)\phi_k(t)dt = \delta_{jk}$$

   其中 δ_{jk} 是 Kronecker Delta（当 $j = k$ 时取 1，否则取 0）。因此，只有当 $j = k$ 时，积分项不为零，所有其他项都消失，得到：

$$\alpha_k \cdot 1 = \int_I X(t)\phi_k(t)dt$$

3. 结论

因此，我们得到了 K-L 展开的系数公式：

$$\alpha_k = \int_I X(t)\phi_k(t)dt$$

这说明 α_k 是随机过程 $X(t)$ 在基函数 $\phi_k(t)$ 方向上的投影（积分形式）。这个公式的意义在于，它可以用来分析和表示随机过程的主要成分，并且对于信号处理、模式识别和数据降维（如 PCA）具有重要应用。

直观解释

如果你熟悉**傅里叶级数展开**，这个公式的作用与傅里叶级数的系数计算类似：

$$c_n = \int X(t)e^{-jn\omega t}dt$$

它表示的是在正交基 $e^{-jn\omega t}$ 上的投影。而 K-L 展开则是用 **一组适合 $X(t)$ 的最佳正交基** 来进行展开，因此有类似的投影系数计算方式。

总结

α_k 可以用 $X(t)\phi_k(t)$ 的积分形式表达，因为：

1. **K-L 展开是一种基函数展开**，其中 $\phi_k(t)$ 是一组正交基。
2. **利用正交性**，投影计算中所有非相关项都消失，只留下对应的系数。
3. 这个积分形式 **本质上是计算 $X(t)$ 在基函数 $\phi_k(t)$ 方向上的投影。**

$I=[-\frac{T}{2},\frac{T}{2}]$

$$\int_{-\frac{T}{2}}^{\frac{T}{2}} R_X(t-s)\phi_k(s)\,ds=\boxed{\lambda_k}\,\phi_k(t)$$

$\phi_k(t)=\exp(i\frac{2k\pi}{T}t)$

$$\int_{-\frac{T}{2}}^{\frac{T}{2}} R_X(\underbrace{t-s}_{s'})\exp(i\frac{2k\pi}{T}\underbrace{s}_{t-s'})\,ds$$

$ds'=-ds$

$s'\in[t-\frac{T}{2},t+\frac{T}{2}]$

$$=\boxed{-\int_{t-\frac{T}{2}}^{t+\frac{T}{2}} R_X(s')\exp(-i\frac{2k\pi}{T}s')\,ds'}\exp(i\frac{2k\pi}{T}t)$$

$R_X(\tau)=R_X(\tau+T)$

① W.S.S. $R_X(\tau)=R_X(\tau+T)$

$$X(t)=\sum_{k=-\infty}^{+\infty}\underline{\alpha_k}\exp(i\frac{2k\pi}{T}t)$$

(双正交 / Bi_Orthogonal)

② $T\to\infty$

$$X(t)\neq\int_{-\infty}^{+\infty}\underline{\alpha(w)}\exp(iwt)\,dw$$

形式化操作

$$=\int_{-\infty}^{+\infty}\exp(iwt)\,d\underline{F_X(w)}$$

Stieltjes Integration	Riemam
$\int f(x)dg(x)\Leftarrow\sum_k f(x_k)\Delta g(x_k)$ $\Delta g(x_k)=g(x_k)-g(x_{k-1})$	$\int f(x)dx\Leftarrow\sum_k f(x_k)\Delta x_k$

$$\boxed{X(t)=\int_{-\infty}^{+\infty}\exp(iwt)\,dF_X(w)}$$

谱函数

↳对应 α_k(双正交)

随机过程的谱表示

Spectral Representation

$$E\left(dF_X(w_1)\overline{dF_X(w_2)}\right)=0\quad (w_1\neq w_2)$$

$\Downarrow$ $w_1=w_2$

$$E|dF_X(w)|^2=\frac{1}{2\pi}S_X(w)dw$$

$$R_X(t,s)=R_X(t-s)=E(X(t)\overline{X(s)})$$

$$=E\left(\int_{-\infty}^{+\infty}\exp(iw_1t)\,dF_X(w_1)\overline{\int_{-\infty}^{+\infty}\exp(iw_2t)\,dF_X(w_2)}\right)$$

$$=\iint\exp(i(w_1t-w_2t))E(dF_X(w_1)\overline{dF_X(w_2)})$$

$$=\int\exp(iw(t-s))\boxed{E|dF_X(w)|^2}$$

$$=\boxed{\frac{1}{2\pi}}\int\exp(iw(t-s))\boxed{S_X(w)dw}$$

(功率谱密度傅立叶反变换)