

研究生教材(试用)

现代数字信号处理

姚天任 孙 洪

华中理工大学电信系

目 录

、数字

第一章 基础知识	1
§ 1.1 随机矢量	1
§ 1.2 相关抵消	3
§ 1.3 Gram-Schmidt 正交化	5
1.3.1 基本定义	5
1.3.2 正交投影定理和 Gram-Schmidt 正交化	8
1.3.3 新息	10
§ 1.4 偏相关系数(PARCOR 系数)	11
§ 1.5 功率谱和周期图	13
§ 1.6 谱分解	15
1.6.1 最小相位序列	15
1.6.2 部分能量和最小延时	16
1.6.3 自相关函数的不变性	17
1.6.4 最小时延性质	17
1.6.5 最小相位性质	18
1.6.6 谱分解定理	19
§ 1.7 信号的参数模型	22
习题	24
参考文献	26
第二章 维纳滤波器和卡尔曼滤波器	27
§ 2.1 维纳滤波问题	27
§ 2.2 维纳滤波器的正交方程和标准方程	29
§ 2.3 FIR 维纳滤波器	31
§ 2.4 IIR 维纳滤波器	33
§ 2.5 维纳滤波器的均方误差	35
§ 2.6 标量卡尔曼滤波器	36
2.6.1 因果 IIR 维纳解	36
2.6.2 维纳解的递推计算—卡尔曼滤波	39
2.6.3 标量卡尔曼滤波器的推导	40
§ 2.7 矢量卡尔曼滤波器	43

复习思考题

习题

§2.8 维纳滤波和卡尔曼滤波的计算和应用

复习思考题

· 2 ·

目 录

2.7.1 信号矢量和数据矢量	43
2.7.2 矢量卡尔曼滤波器的推导	45
习题	49
参考文献	52
第三章 自适应滤波器	53
§ 3.1 自适应滤波原理	53
§ 3.2 自适应线性组合器	55
§ 3.3 均方误差性能曲面	57
§ 3.4 二次性能曲面的基本性质	59
§ 3.5 最陡下降法	62
§ 3.6 学习曲线和收敛速度	65
§ 3.7 最小均方(LMS)自适应算法	69
§ 3.8 权向量噪声	71
§ 3.9 失调量	73
§ 3.10 递归最小二乘方(RLS)自适应算法	75
§ 3.11 无限冲激响应(IIR)自适应滤波器	80
3.11.1 IIR 自适应滤波器的 LMS 算法	81
3.11.2 超稳定自适应递归滤波器(SHARP 算法)	83
§ 3.14 自适应滤波器的应用	83
3.14.1 自适应系统模拟和辨识	83
3.14.2 自适应逆滤波或逆向模拟	85
3.14.3 自适应干扰抵消	86
3.14.4 自适应预测	88
习题	88
参考文献	91
第四章 功率谱估计的现代方法	92
§ 4.1 从经典谱估计到现代谱估计	92
§ 4.2 谱估计的参数模型方法	94
§ 4.3 AR 模型的 Yule-Walker 方程	97
§ 4.4 Levinson-Durbin 算法	98
§ 4.5 AR 模型的稳定性和它的阶的确定	104
§ 4.6 AR 谱估计的性质	108
4.6.1 AR 谱估计隐含着自相关函数的外推	108
4.6.2 AR 谱估计与一维高斯过程最大熵谱估计等效	109

§3.12 自适应滤波器计算举例

§3.13 自适应滤波器的数字实现

3.13.1 LMS 算法自适应滤波器的直接实现

3.13.2 分布运算自适应滤波器

3.13.3 系数制自适应滤波器

复习思考题

4.11.1 最大似然法	4.6.3 AR谱估计与线性预测谱估计等效	110
	4.6.4 AR谱估计可看成最佳白化处理结果	111
4.11.2 修正协方差AR谱估计方法	4.6.5 AR谱估计的界	114
	§ 4.7 格形滤波器	114
4.11.3 特征分解频率估计	§ 4.8 提取AR模型参数的方法	117
	4.8.1 Yue-Walker或自相关法	118
	4.8.2 协方差法	119
	4.8.3 Burg法	120
复习思考题	§ 4.9 AR谱估计的异常现象及其补救措施	123
	4.9.1 虚假谱峰	123
	4.9.2 谱线分裂	123
	4.9.3 噪声对AR谱估计的影响	124
附录4.1~附录4.11	习题	127
	参考文献	129
	第五章 同态信号处理	130
	§ 5.1 广义迭加原理	130
	§ 5.2 乘法同态系统	132
	§ 5.3 卷积同态系统	135
	§ 5.4 复倒谱定义	139
	5.4.1 复对数的多值性问题	139
	5.4.2 $\hat{Z}(a)$ 的解析问题	140
	§ 5.5 复倒谱的性质	141
	§ 5.6 复倒谱的计算方法	142
	5.6.1 按复倒谱定义计算	142
	5.6.2 最小相位序列的复倒谱的计算	146
	5.6.3 复对数求导数算法	148
	5.6.4 递推计算方法	150
	习题	152
	参考文献	154

第一章 基础知识

本章主要目的是复习本书以下各章将要用到的有关随机过程和估计理论的部分内容，也起着统一全书符号的作用。

§ 1.1 随机矢量

离散随机信号或过程，是由随机变量按一定顺序排列而成的时间序列。将其看成随机矢量有时是方便的。即离散随机信号可表示为：

$$\mathbf{X} = (x_0, x_1, \dots, x_n, \dots)^T \quad (1.1)$$

它可以是有限维的也可以是无限维的。

若随机信号 x_n 的均值 $E[x_n] \neq 0$ ，可将其移去，即重新定义一个零均值随机信号 $x_n - E[x_n]$ ，因此本书今后将只讨论零均值随机信号。

最简单的随机信号是零阶马尔柯夫信号或纯随机信号，它的所有随机变量是互相独立的，且有共同的密度函数 $P(x)$ ，因此有

$$P(x_0, x_1, \dots, x_n, \dots) = P(x_0)P(x_1)\cdots P(x_n)\cdots \quad (1.2)$$

$$R_{xx}(n, m) = E[x_n x_m] = 0, \quad n \neq m$$

这里， $P(x_0, x_1, \dots, x_n, \dots)$ 是联合概率密度函数， $R_{xx}(n, m)$ 是 x_n 和 x_m 的相关函数。这种随机信号是平稳的，其自相关函数与时间起点无关，且各随机变量的方差相同，即

$$E[x_n^2] = R_{xx}(0) = \sigma_x^2 \quad (1.3)$$

这样，纯随机信号的自相关函数可表示为

$$R_{xx}(k) = E[x_{n+k} x_n] = \sigma_x^2 \delta(k) \quad (1.4)$$

这里 $\delta(k)$ 是单位取样序列。

纯随机信号无记忆，因为

$$P(x_n, x_{n-1}) = P(x_n)P(x_{n-1})$$

或

$$P(x_n / x_{n-1}) = P(x_n) \quad (1.5)$$

这意味着， x_{n-1} 的发生不影响 x_n 的数值。由于一个随机过程中相继的随机变量的取值完全是随机的，因此这种信号表现出随着时间有很快的变化，但它同时也会表现出随着时间的慢变化。这样，它含有从很低到很高的几乎所有频率，而且比例相同。显然这就是白噪声。

一阶马尔柯夫信号稍微复杂一些，它有一个取样时间间隔的记忆能力，即

$$P(x_n / x_{n-1}, x_{n-2}, \dots, x_1, x_0) = P(x_n / x_{n-1}) \quad (1.6)$$

这就是说， x_n 只受前一个取样值 x_{n-1} 的影响。一阶马尔柯夫信号的完全统计描述很简单，只需边界密度函数 $P(x_n)$ 和条件密度函数 $P(x_n / x_{n-1})$ 足够。任何其它联合概率密度函数都可由此计算出来。

一个 N 维随机矢量 X 是由 N 个不同随机变量为分量构成的，即

$$X = (x_1, x_2, \dots, x_N)^T \quad (1.7)$$

各分量之间可以有相依性也可以没有。为了完全统计描述一个 N 维随机矢量，需要知道联合概率密度函数

$$P(X) = P(x_1, x_2, \dots, x_N) \quad (1.8)$$

N 维随机矢量 X 的均值是一个 N 维矢量

$$m = E[X] \quad (1.9)$$

X 的自相关函数是一个 $N \times N$ 的正半定对称矩阵

$$R = E[XX^T] \quad (1.10)$$

X 的自协方差函数也是一个 $N \times N$ 的正半定对称矩阵

$$\Sigma = E[(X-m)(X-m)^T] \quad (1.11)$$

以上各式中的期望运算都是用联合概率密度函数来定义的，例如

$$E[X] = \int X P(X) d^N X \quad (1.12)$$

这里, $P(X)$ 是由式 (1.8) 所定义的联合概率密度函数; 而 $d^N X = dx_1 dx_2 \cdots dx_N$; 上标 T 表示矩阵转置.

自相关矩阵和自协方差矩阵之间有下列关系:

$$\Sigma = R - \mu \mu^T \quad (1.13)$$

在均值为零情况下, Σ 与 R 相等.

§ 1.2 相关抵消

设 X 和 Y 分别是 N 维和 M 维零均值随机矢量, 且它们相关, 即有

$$R_{xy} = E[XY^T] \neq 0 \quad (1.14)$$

现对 Y 进行线性变换, 得

$$\hat{X} = HY \quad (1.15)$$

这里 H 是 $N \times M$ 变换矩阵. 适当选择 H , 使随机矢量

$$e = X - \hat{X} = X - HY \quad (1.16)$$

与 Y 不相关, 即

$$E_{ey} = E[eY^T] = 0 \quad (1.17)$$

将式(1.16)代入式(1.17), 得

$$R_{ey} = E[XY^T] - HE[YY^T] = R_{xy} - HR_{yy} = 0$$

由此求得

$$H = R_{xy} R_{yy}^{-1} = E[xy^T] E[yy^T]^{-1} \quad (1.18)$$

这就是说, 若按式 (1.18) 选择线性变换矩阵 H , 则 X 中与 Y 相关的部分即 $\hat{X} = HY$ 将被消除. 图 1.1 说明这一处理过程, 并称为相关抵消器. 它具有以下三个功能:

(1) 最佳线性估计

设 $X = X_1 + X_2$, X_1 与 Y 相关, X_2 与 Y 不相关. Y 经线性变换后得到 X_1 的估计 $\hat{X} = HY$, 估计误差为 $e_1 = X_1 - \hat{X} = X_1 - Hy$. 在均方误差最小的意义上这种估计是最佳的, 即 H 的选择准则是使估计误差的均方值最小

$$R_{e_1 e_1} = E[e_1 e_1^T] = \min \tag{1.19}$$

$R_{e_1 e_1}$ 对 H 求导数, 并令其等于零, 即可解出最佳线性估计所要求的 H 值

$$H_{opt} = R_{xy} R_{yy}^{-1} \tag{1.20}$$

这里 H 的下标表示“最佳”. 该式与式 (1.18) 等效.

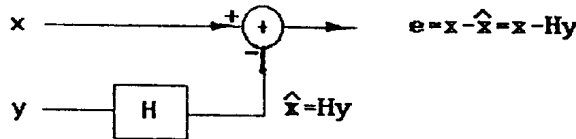


图 1.1 相关抵消器原理图

(2) 相关抵消

输出 $e = X - \hat{X} = X_2 + (X_1 - \hat{X})$ 是 X_2 与估计误差 $(X_1 - \hat{X})$ 之和. X 中与 Y 相关的部分即 X_1 已被抵消, 若 H 是最佳估计, 则抵消后剩余部分即 $X_1 - \hat{X}$ 具有最小均方值.

(3) 最佳信号分离

由于 \hat{X} 是 X_1 的最佳估计即 $\hat{X} \approx X_1$, 故有 $e \approx x_2$. 这样, 相关抵消器的两个输出分别是 X_1 和 X_2 的近似值, 具有分离 X 中两部分的效果.

许多实际应用中, 线性变换矩阵 H 常表示一种线性滤波运算, 矢量 X 和 Y 是信号取样序列. 由式 (1.18) 看出, 为了设计 H , 需要知道 R_{xy} 和 R_{yy} . 常有两种情况. 一种情况是已知 X 和 Y , 要求消除它们之间的相关性. 若用时间平均代替集合平均或采用自适应技术, 即可利用 X 和 Y 的数据来估计 R_{xy} 和 R_{yy} . 这种情况的典型应用有噪声抵消、回波抵消、信道均衡、天线旁瓣抵消等. 另一种情况是只知道 Y , 并不知道 X , 需要根据 Y 来估计 X . 为此必需首先建立一个模型来描述 Y 与 X 之间的关系, 然后据此模型计算 R_{xy} 和 R_{yy} . 这种情况的典型应用是卡尔曼滤波器. 以上两种情况都将在本书中详细讨论.

§ 1.3 Gram-Schmidt 正交化

从前节得知, 任一随机矢量 \mathbf{X} 相对于另一矢量 \mathbf{Y} 可分解为两部分, 即 $\mathbf{X} = \hat{\mathbf{X}} + \mathbf{e}$, 其中一部分 $\hat{\mathbf{X}}$ 与 \mathbf{Y} 相关即 $\hat{\mathbf{X}} = \mathbf{H}\mathbf{Y}$, 另一部分 \mathbf{e} 与 \mathbf{Y} 不相关即 $R_{\mathbf{e}\mathbf{Y}} = 0$. 这两部分之间的相关函数为

$$R_{\mathbf{e}\hat{\mathbf{X}}} = E[\mathbf{e}\hat{\mathbf{X}}^T] = E[\mathbf{e}\mathbf{Y}^T\mathbf{H}^T] = R_{\mathbf{e}\mathbf{Y}}\mathbf{H}^T = 0 \quad (1.21)$$

因而它们是不相关的。

本节将要进一步做几何解释, 说明它们还是互相正交的。几何解释有两方面意义: 第一, 为推导和理解信号估计问题提供了一种简单和直观的方法; 第二, 通过 Gram-Schmidt 正交化方法, 为建立信号模型奠定了基础, 信号模型在语音处理、数据压缩和现代谱估计中有着广泛应用。

1.3.1 基本定义

设 \mathbf{X} 和 \mathbf{Y} 分别是 N 维和 M 维随机矢量

$$\mathbf{X} = (x_1, x_2, \dots, x_N)$$

$$\mathbf{Y} = (y_1, y_2, \dots, y_M)$$

所要讨论的随机变量空间是下列随机变量集合生成的 $N+M$ 维线性矢量空间

$$\{x_1, x_2, \dots, x_N, y_1, y_2, \dots, y_M\}$$

设 \mathbf{u} 和 \mathbf{v} 是该线性空间中任二矢量, 并定义它们的内积为

$$(\mathbf{u}, \mathbf{v}) = E(\mathbf{u}^T\mathbf{v}) \quad (1.22)$$

该线性空间是一个内积空间。内积空间中两矢量 \mathbf{u} 和 \mathbf{v} 间的距离用内积定义为

$$|\mathbf{u} - \mathbf{v}| = \sqrt{(\mathbf{u} - \mathbf{v}, \mathbf{u} - \mathbf{v})} = \sqrt{E[(\mathbf{u} - \mathbf{v})^2]} \quad (1.23)$$

非零矢量 \mathbf{u} 和 \mathbf{v} 的夹角 θ 的余弦规定为

$$\cos\theta = \frac{(\mathbf{u}, \mathbf{v})}{|\mathbf{u}||\mathbf{v}|} \quad (1.24)$$

假如 $(\mathbf{u}, \mathbf{v}) = 0$, 那末 $\cos\theta = 0$, 这时 \mathbf{u} 和 \mathbf{v} 叫做正交, 记成 $\mathbf{u} \perp \mathbf{v}$ 。

在内积空间中一组线性无关的矢量不一定两两正交, 但是一组两两正交的非零矢量是

线性无关的。设有 M 个两两正交的随机矢量 $\epsilon_1, \epsilon_2, \dots, \epsilon_M$ ，它们满足下式

$$(\epsilon_i, \epsilon_j) = 0, \quad i \neq j \quad (1.25)$$

令 $Y = \{\epsilon_1, \epsilon_2, \dots, \epsilon_M\}$ 是它们张成的线性子空间，那么可用它们作为该内积空间的正交基底。

根据正交分解定理：任何随机变量 x ，关于线性子空间 Y ，可唯一分解为两互相正交的部分，一部分位于 Y 内（与 Y 平行），另一部分与 Y 垂直，即

$$x = \hat{x} + e, \quad \hat{x} \in Y, \quad e \perp Y \quad (1.26)$$

分量 \hat{x} 称为 x 在 Y 上的正交投影。图 1.2 举例说明正交分解定理的几何意义，这里假设子空间 Y 是二维空间（平面），随机变量 x 和它的两个分量 \hat{x} 和 e 都用矢量表示。在式 (1.26) 中，正交条件 $e \perp Y$ 意味着 e 与 Y 中每个矢量正交，当然也与 Y 的每个基底矢量 ϵ_i 正交，即有

$$(e, \epsilon_i) = E[ee_i] = 0 \quad (1.27)$$

该式称为正交方程。

\hat{x} 位于 Y 内，因此可用正交基底的线性组合来表示

$$\hat{x} = \sum_{i=1}^M a_i \epsilon_i \quad (1.28)$$

系数 a_i 可根据正交方程确定。因为

$$(x, \epsilon_i) = (\hat{x} + e, \epsilon_i) = (\hat{x}, \epsilon_i) + (e, \epsilon_i)$$

利用正交方程，上式简化为

$$(x, \epsilon_i) = (\hat{x}, \epsilon_i)$$

将式(1.28)代入上式，得

$$(x, \epsilon_i) = \sum_{j=1}^M a_j (\epsilon_j, \epsilon_i)$$

由于 ϵ_i 互相正交(式(1.25))，上式遂变为

$$(x, \epsilon_i) = a_i (\epsilon_i, \epsilon_i)$$

由此求得

$$a_i = (x, \epsilon_i)(\epsilon_i, \epsilon_i)^{-1} = E[x\epsilon_i]E[\epsilon_i\epsilon_i]^{-1} \quad (1.29)$$

这就是计算 a_i 的公式。

将式(1.29)和式(1.28)代入式(1.26), 得到

$$x = \sum_{i=1}^M E[x\epsilon_i]E[\epsilon_i\epsilon_i]^{-1}\epsilon_i + e \quad (1.30)$$

写成矩阵形式

$$x = E[x\epsilon]E[\epsilon\epsilon^T]^{-1}\epsilon + e \quad (1.31)$$

式中,

$$\begin{cases} \epsilon = (\epsilon_1, \epsilon_2, \dots, \epsilon_M)^T \\ E[x\epsilon] = [E(x\epsilon_1), E(x\epsilon_2), \dots, E(x\epsilon_M)]^T \\ E[\epsilon\epsilon^T] = \text{diag} \{E(\epsilon_1^2), E(\epsilon_2^2), \dots, E(\epsilon_M^2)\} \end{cases} \quad (1.32)$$

符号 diag 表示对角矩阵。

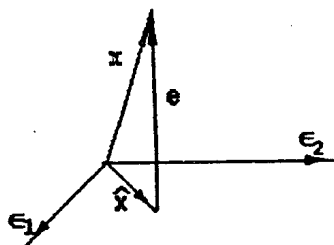


图 1.2 正交分解定理的几何意义

利用式(1.32)规定的符号, 正交方程的矩阵形式为

$$E_{ee} = E[ee^T] = 0 \quad (1.33)$$

(1.31)和式(1.32)是随机变量 x 关于随机变量线性子空间 Y 的唯一正交分解的表示式。

如果将 N 个随机变量 x_1, x_2, \dots, x_N 关于子空间 Y 进行正交分解, 那么有

$$x_i = \hat{x}_i + e_i, \quad i = 1, 2, \dots, N$$

这 N 个方程可用矢量矩阵形式合写成

$$\mathbf{X} = \hat{\mathbf{X}} + \mathbf{e} = \mathbf{E}[\mathbf{X}\mathbf{e}^T] \mathbf{E}[\mathbf{e}\mathbf{e}^T]^{-1} \mathbf{e} + \mathbf{e} \quad (1.34)$$

正交方程可表示为

$$\mathbf{R}_{\mathbf{e}\mathbf{e}} = \mathbf{E}[\mathbf{e}\mathbf{e}^T] = 0 \quad (1.35)$$

式(1.34)中, $\mathbf{X} = (x_1, x_2, \dots, x_N)^T$ 是 N 维列矢量.

1.3.2 正交投影定理和 Gram-Schmidt 正交化

正交投影定理: 矢量 x 在线性子空间 Y 上的正交投影 \hat{x} 是 Y 中与 x 距离最近的一个矢量.

证明:

根据正交分解定理, x 关于 Y 的唯一正交分解表示为

$$x = \hat{x} + e, \quad \hat{x} \in Y, \quad e \perp Y \quad (1.36)$$

设 y 是 Y 中任一矢量, 由于 $(\hat{x} - y) \in Y$, 因而 $e \perp (\hat{x} - y)$, 故有

$$|x - y|^2 = |(\hat{x} - y) + e|^2 = |\hat{x} - y|^2 + |e|^2$$

由式(1.23), 上式可写成

$$\mathbf{E}[(x - y)^2] = \mathbf{E}[|\hat{x} - y|^2 + |e|^2] = \mathbf{E}[|\hat{x} - y|^2] + \mathbf{E}[|e|^2] \quad (1.37)$$

由该式看出, 当 $y = \hat{x}$ 时, x 与 y 间距离最近. 定理得证.

正交投影定理说明: 用 Y 中随机变量的线性组合来逼近 x 时, 在最小二乘方的意义上, \hat{x} 是最佳的.

实际中, 线性子空间 Y 常由一组非正交基底定义, 设它们是

$$Y = \{y_1, y_2, \dots, y_n\}$$

式中各随机变量可以是相关的.

Gram-Schmidt 正交化是一个递归处理过程, 其目的是由非正交基底 $\{y_1, y_2, \dots, y_M\}$ 求出一组正交基底 $\{\epsilon_1, \epsilon_2, \dots, \epsilon_n\}$. 基本思路是: 先选择 y_1 作为正交基底 ϵ_1 ; 其次, 将 y_2 关于 ϵ_1 进行正交分解并选择与 ϵ_1 垂直的分量作为 ϵ_2 , 于是有 $(\epsilon_1, \epsilon_2$

2) = 0; 再次, 将 y_3 关于子空间 $\{\epsilon_1, \epsilon_2\}$ 的分量作为 ϵ_3 , 类此不断继续做下去, 整个处理过程表示如下:

$$\begin{aligned}
 \epsilon_1 &= y_1 \\
 \epsilon_2 &= y_2 - E[y_2 \epsilon_1] E[\epsilon_1 \epsilon_1]^{-1} \epsilon_1 \\
 \epsilon_3 &= y_3 - [E[y_3 \epsilon_1] E[\epsilon_1 \epsilon_1]^{-1} \epsilon_1 + E[y_3 \epsilon_2] E[\epsilon_2 \epsilon_2]^{-1} \epsilon_2] \\
 &\vdots \\
 &\vdots \\
 &\vdots \\
 \epsilon_n &= y_n - \sum_{i=1}^{n-1} E[y_n \epsilon_i] E[\epsilon_i \epsilon_i]^{-1} \epsilon_i, \quad 2 \leq n \leq m
 \end{aligned} \tag{1.38}$$

这样构造出来的基底 $\{\epsilon_1, \epsilon_2, \dots, \epsilon_m\}$ 就是 Y 的正交基底.

Gram-Schmidt 正交化过程, 实际上可理解为由前一子空间增加一个正交基底以得到后一子空间, 从而不断扩大子空间的过程, 如下式所示:

$$\begin{cases}
 Y_1 = \{\epsilon_1\} \\
 Y_2 = \{\epsilon_1, \epsilon_2\} \\
 Y_3 = \{\epsilon_1, \epsilon_2, \epsilon_3\} \\
 \vdots \\
 Y_n = \{\epsilon_1, \epsilon_2, \dots, \epsilon_n\}
 \end{cases} \tag{1.39}$$

式(1.38)右边第二项是 y_n 在子空间 Y_{n-1} 上的正交投影, 用符号 $\hat{y}_{n|n-1}$ 来表示, 即

$$\hat{y}_{n|n-1} = \sum_{i=1}^{n-1} E[y_n \epsilon_i] E[\epsilon_i \epsilon_i]^{-1} \epsilon_i \tag{1.40}$$

于是式(1.38)可写成

$$\epsilon_n = y_n - \hat{y}_{n|n-1} \text{ 或 } y_n = \hat{y}_{n|n-1} + \epsilon_n \tag{1.41}$$

实际上这就是 y_n 在子空间 Y_{n-1} 上的正交分解.

引用符号

$$\begin{aligned}
 b_{ni} &= E[y_n \epsilon_i] E[\epsilon_i \epsilon_i]^{-1}, \quad 1 < i < n-1 \\
 b_{nn} &= 1
 \end{aligned}$$

式(1.41)可写成

$$y_n = \sum_{i=1}^n b_{ni} \epsilon_i, \quad 1 \leq n \leq M \quad (1.42)$$

上式写成矩阵形式

$$Y = B\epsilon \quad (1.43)$$

这里, $Y = (y_1, \dots, y_M)^T$

$$\epsilon = (\epsilon_1, \epsilon_2, \dots, \epsilon_M)^T$$

$$B = \begin{bmatrix} 1 & 0 & 0 & 0 & \cdots & 0 \\ b_{21} & 1 & 0 & 0 & \cdots & 0 \\ b_{31} & b_{32} & 1 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ b_{M1} & b_{M2} & b_{M3} & b_{M4} & \cdots & 1 \end{bmatrix} \quad (1.44)$$

1.3.3 新息

两组基底 Y 和 ϵ 可以构成同一矢量空间 Y , 它们所含信息相同. 正交化结果, 基底 y 变成基底 ϵ , 只是去掉了 y 中元素之间的相关性或冗余信息. 基底 ϵ 中各分量不相关, 分别包含着各自不同的信息, 因此每增加一个基底 ϵ_i 意味着增加新的信息. 常将随机变量称为新息, 而式(1.43)称为 y 的新息表示.

由式 (1.43), 计算 Y 的相关矩阵, 得

$$R_{yy} = E[yy^T] = BR_{\epsilon\epsilon}B^T \quad (1.45)$$

式中, $R_{\epsilon\epsilon} = E[\epsilon\epsilon^T]$ 是对角矩阵, B 如式 (1.44) 所示是下三角矩阵, B^T 是 B 的转置为上三角矩阵. 式 (1.45) 实际上是 R_{yy} 矩阵的 Cholesky 分解.

在 1.3.1 节中讨论过, 随机变量 x 相对于线性子空间 Y 进行正交分解, 得到式 (1.31) 的结果, 其中

$$\hat{x} = E[x\epsilon^T]E[\epsilon\epsilon^T]^{-1}\epsilon$$

是 x 在 Y 上的正交投影, 这里 ϵ 是 Y 的正交基底. 将式 (1.43) 代入上式

$$\begin{aligned} \hat{x} &= E[x\epsilon^T]E[\epsilon\epsilon^T]^{-1}\epsilon \\ &= E[xY^T(B^T)^{-1}][B^{-1}yy^TB]B^{-1}y \\ &= E[xY^T]E[yy^T]^{-1}Y \end{aligned} \quad (1.46)$$

这说明, Y 的基底变化后, x 在 Y 上的正交投影并未发生改变.

利用式 (1.46), 将 y_n 在子空间 $Y_{n-1} = \{y_1, y_2, \dots, y_{n-1}\}$ 上进行正交分解, 正交投影为

$$\hat{y}_{n|n-1} = E[y_n Y_{n-1}^T] E[Y_{n-1} Y_{n-1}^T]^{-1} Y_{n-1}$$

正交分量为

$$\epsilon_n = Y_n - \hat{Y}_{n|n-1} = Y_n - E[y_n Y_{n-1}^T] E[Y_{n-1} Y_{n-1}^T]^{-1} Y_{n-1} \quad (1.47)$$

该式提供了直接由 y_n 计算 ϵ_n 的简捷方法.

$\hat{y}_{n|n-1}$ 是根据 $Y_{n-1} = (y_1, y_2, \dots, y_{n-1})^T$ 对 y_n 做出的最佳线性估计. 若下标 n 表示时间, $\hat{y}_{n|n-1}$ 便是过去值对 y_n 作出的最佳线性预测, 而 ϵ_n 便是相应的预测误差.

预测误差或新息序列具有一些有用的性质, 将在第二章加以介绍.

§ 1.4 偏相关系数 (PARCOR 系数)

设 N 维随机矢量 Y 被任意分成三个矢量,

$$Y = (y_0, y_1, \dots, y_N)^T = (Y_0^T, Y_1^T, Y_2^T)^T \quad (1.48)$$

如果各随机变量 y_i 是相关的, 那么一般而言, Y_0, Y_1 和 Y_2 这三个矢量也应当是两两相关的. 现来讨论 Y_0 和 Y_2 之间的相关性. 这种相关性可区分为两部分, 一部分称为间接相关, 它的产生是由于 Y_0 与 Y_1 相关而 Y_1 又与 Y_2 相关, 所以才有 Y_0 与 Y_2 之间的相关; 另一部分称为直接相关, 它与 Y_1 存在与否没有关系, 或者说, 即使 Y_1 不存在, Y_0 与 Y_2 之间也存在着这种直接相关性, 称为 Y_0 与 Y_2 之间的偏相关性, 其大小用偏相关系数或 PARCOR 系数来度量.

图 1.3 示出 Y_0 关于 Y_1 和 Y_2 关于 Y_1 的正交分解情况. Y_0 与 Y_2 的偏相关系数定义为它们关于 Y_1 分解的正交分量 e_0 和 e_2 间的相关系数的归一化值, 归一化系数为 $E[e_0 e_2^T]$.

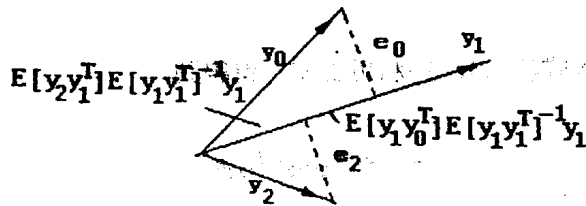


图 1.3 Y_0 和 Y_2 关于 Y_1 的正交分解

根据正交分解定理, 有

$$Y_0 = e_0 + E[Y_0 Y_1^T] E[Y_1 Y_1^T] Y_1 \quad (1.49)$$

$$Y_2 = e_2 + E[Y_2 Y_1^T] E[Y_1 Y_1^T] Y_1 \quad (1.50)$$

将 $Y = (Y_0^T, Y_1^T, Y_2^T)^T$ 正交化得到 $\epsilon = (\epsilon_0^T, \epsilon_1^T, \epsilon_2^T)^T$, 即

$$T = B\epsilon$$

或

$$\epsilon = B^{-1}Y = AY \quad (1.50)$$

这里

$$A = B^{-1}$$

将式(1.50)写成矩阵形式,

$$\begin{bmatrix} \epsilon_0 \\ \epsilon_1 \\ \epsilon_2 \end{bmatrix} = \begin{bmatrix} A_{00} & 0 & 0 \\ A_{11} & A_{10} & 0 \\ A_{22} & A_{21} & A_{20} \end{bmatrix} \begin{bmatrix} Y_0 \\ Y_1 \\ Y_2 \end{bmatrix} \quad (1.51)$$

由上式得出

$$\epsilon_2 = A_{22}Y_0 + A_{21}Y_1 + A_{20}Y_2 \quad (1.52)$$

因 $\epsilon_2 \perp Y_1$, 故有

$$E[\epsilon_2 Y_1^T] = A_{22}E[Y_0 Y_1^T] + A_{21}E[Y_1 Y_1^T] + A_{20}E[Y_2 Y_1^T] = 0$$

将式(1.49)和式(1.50)代入式(1.52), 并利用上式, 便得到

$$\epsilon_2 = A_{22}e_0 + A_{20}e_2 \quad (1.53)$$

因 $e_0 \in \{Y_0^T, Y_1^T\}$, $\epsilon_2 \perp [Y_0^T, Y_1^T]$, 故 $\epsilon_2 \perp e_0$, 即 $E[\epsilon_2 e_0^T] = 0$, 由式(1.53)得

$$A_{22}E[\mathbf{e}_0\mathbf{e}_0^T] + A_{20}E[\mathbf{e}_2\mathbf{e}_0^T] = 0$$

由该式求出

$$\Gamma = E[\mathbf{e}_2\mathbf{e}_0^T]E[\mathbf{e}_0\mathbf{e}_0^T]^{-1} = -A_{20}^{-1}A_{22} \quad (1.54)$$

偏相关是一个与 Gram-Schmidt 正交化紧密相关的概念，它在线性预测和现代谱估计中起着很重要的作用。

§ 1.5 功率谱和周期图

离散时间实平稳随机信号 x_n 的功率谱 $S_{xx}(z)$ 定义为它的自相关函数 $R_{xx}(k)$ 的双边 z 变换，

$$S_{xx}(z) = \sum_{k=-\infty}^{\infty} R_{xx}(k)z^{-k} \quad (1.55)$$

式中， $R_{xx}(k)$ 定义为

$$R_{xx}(k) = E[x_{n+k}x_n] \quad (1.56)$$

若 $R_{xx}(k)$ 是稳定的，则 $S_{xx}(z)$ 的收敛域包括 z 平面上的单位圆，于是令 $z = e^{j\omega}$ 便可由式 (1.55) 定义以频率为自变量的功率谱，

$$S_{xx}(\omega) = \sum_{k=-\infty}^{\infty} R_{xx}(k)z^{-jk} \quad (1.57)$$

$S_{xx}(\omega)$ 在物理意义上说明了信号 x_n 的频率成分，以及功率随频率的分布。不相关随机信号（白噪声）的自相关函数是单位取样序列，如式 (1.4) 所示。其功率谱等于常数 σ_x^2 。

两实平稳随机信号 x_n 和 y_n 的互相关函数和互相率谱分别定义为

$$R_{xy}(k) = E[x_{n+k}y_n] \quad (1.58)$$

$$S_{xy}(z) = \sum_{k=-\infty}^{\infty} R_{xy}(k)z^{-k} \quad (1.59)$$

它们有以下重要性质