北京师范大学现代数学丛书

增长曲线模型中的参数估计

徐承彝　杨文礼　蒋文江　编著

北京师范大学出版社

国家自然科学基金资助项目

增长曲线模型中的参数估计

徐承彝　杨文礼　蒋文江　编著

北京师范大学出版社

图书在版编目(CIP)

增长曲线模型中的参数估计/徐承彝等编著．—北京：北京师范大学出版社，1996．3
ISBN 7-303-04151-6

Ⅰ．增…　Ⅱ．徐…　Ⅲ．线性模型-参数估计　Ⅳ．O211．67

中国版本图书馆 CIP 数据核字(96)第 01218 号

北京师范大学出版社出版发行
(100875　北京新街口外大街 19 号)
北京师范大学印刷厂印刷　全国新华书店经销
开本：850×1168　1/32　印张：8．375　字数：203 千
1996 年 7 月北京第 1 版　1996 年 7 月北京第 1 次印刷
印数：1—2 000 册
定价：11．40 元

前　言

线性模型

$$y = X\beta + \varepsilon$$

是有广泛应用的、重要的统计模型，其中 ε 是 n 维随机变量，数学期望 $E\varepsilon=0$，协方差阵 $E\varepsilon\varepsilon'=\sigma^2 I_n$，$I_n$ 表示 n 阶单位阵，X 是已知的 n 行 k 列矩阵，X 的秩 $\mathrm{rk}X=k<n$，k 维向量 β 是未知参数向量，$\sigma^2>0$ 是未知参数。

通常使用的 σ^2 的估计量是

$$\hat{\sigma}^2 = \frac{1}{n-k} y' [I_n - X(X'X)^{-1}X'] y,$$

这是一个 y 的二次型，是 σ^2 的无偏估计量。

对于无偏估计量，自然要研究它有没有最小方差性质。$\hat{\sigma}^2$ 在 σ^2 的二次型无偏估计类中的最小方差性质，最早的结果是我国许宝騄教授在 1938 年得到的，这就是著名的许定理[1]。许定理给出了 $\hat{\sigma}^2$ 是 σ^2 的最小方差不变二次无偏估计的充要条件，这项研究在数理统计中是一项奠基性的工作。由于定理的证明比较困难，在以后的几十年中，随着数学工具的发展，不断地出现新的证明方法，例如[2]、[3]、[4]，也有一些文章给出了一些充分条件，例如[5]、[6]。在许定理发表 40 年之后，1979 年 J・Kleffe 把许定理推广到多元线性模型中，得到了类似的结果——多元许定理[7]。

我们及我们指导的研究生在更为一般的多元线性模型中推广了 Kleffe 的工作，对二次型估计的一些优良性质进行了更为深入的研究。本书将首先介绍许定理及多元许定理，然后，在此基础上介绍我们的工作。

张尧庭教授一直关心和鼓励我们的工作，这次又仔细地审阅了本书手稿，提出了很多修改意见，对此我们表示衷心的感谢。北京师范大学出版社对本书的出版给予了热心的支持，在此也表示衷心的感谢。

本书第四章由蒋文江执笔，第三、五章由杨文礼执笔，第一、二、六章由徐承彝执笔。全书由徐承彝统一校阅。由于我们的水平所限，书中会有许多不妥之处，恳请同行专家及读者批评指正。

作者

1994 年 4 月

目　　录

第一章　许定理与多元许定理

1.1　问题与常用符号

考虑线性模型

(1.1.1)　　$$y=X\beta+\varepsilon,$$

其中 ε 是 n 维随机变量，数学期望 $\mathrm{E}\varepsilon=0$，协方差阵 $\mathrm{Cov}\varepsilon=\sigma^2I_n$，$I_n$ 表示 n 阶单位阵，X 是已知的 n 行 k 列矩阵，X 的秩 $\mathrm{rk}X=r<n$，k 维向量 β 是未知的向量参数，$\sigma^2>0$ 是未知参数。

通常使用的 σ^2 的估计量是

(1.1.2)　　$$\hat{\sigma}^2=\frac{1}{n-r}y'[I_n-X(X'X)^-X']y,$$

$\hat{\sigma}^2$ 是 y 的二次型，其中 $(X'X)^-$ 是 $X'X$ 的广义逆矩阵。由于

$$\begin{aligned}(n-r)\mathrm{E}\,\hat{\sigma}^2&=\mathrm{E}y'[I_n-X(X'X)^-X']y\\&=\mathrm{E}\varepsilon'[I_n-X(X'X)^-X']\varepsilon\\&=\mathrm{Etr}\{[I_n-X(X'X)^-X']\varepsilon\varepsilon'\}\\&=\sigma^2\{\mathrm{tr}(I_n)-\mathrm{tr}[X(X'X)^-X']\}\\&=\sigma^2(n-r),\end{aligned}$$

$\hat{\sigma}^2$ 是 σ^2 的无偏估计量。

如果 y 还服从正态分布，$\hat{\sigma}^2$ 还是 σ^2 的最小方差无偏估计。

(1.1.3)　**定理**　若 $y\sim N_n(X\beta,\sigma^2I_n)$，则 $\hat{\sigma}^2$ 是 σ^2 的最小方差无偏估计量。

证　$y\sim N_n(X\beta,\sigma^2I_n)$ 时，$\frac{\varepsilon}{\sigma}\sim N_n(0,I_n)$。

由于 $I_n - X(X'X)^- X'$ 是秩为 $n-r$ 的对称幂等阵，所以

$$\left(\frac{\varepsilon}{\sigma}\right)'[I_n - X(X'X)^- X']\frac{\varepsilon}{\sigma} \sim \chi^2(n-r),$$

$\hat{\sigma}^2$ 的方差

$$\begin{aligned}\mathrm{Var}\,\hat{\sigma}^2 &= \mathrm{Var}\frac{\sigma^2}{n-r}\left(\frac{\varepsilon}{\sigma}\right)'[I_n - X(X'X)^- X']\frac{\varepsilon}{\sigma} \\ &= \frac{\sigma^4}{(n-r)^2}\cdot 2(n-r) \\ &= \frac{2\sigma^4}{n-r}\end{aligned}$$

有限。

设 $X=(x_{ij})$，$\beta=(\beta_1,\cdots,\beta_k)'$，$y=(y_1,\cdots,y_n)'$，$(X'X)^- = (a_{ij})$。若统计量 $T(y)$ 的数学期望为零，方差有限，则有

$$\int\underset{R^n}{\cdots}\int T(y_1,\cdots,y_n)(2\pi\sigma^2)^{-\frac{n}{2}}$$

$$\times\exp\left\{-\frac{1}{2\sigma^2}(y-X\beta)'(y-X\beta)\right\}dy_1\cdots dy_n = 0,$$

或

$$Q(\beta_1,\cdots,\beta_k,\sigma^2) = \int\underset{R^n}{\cdots}\int T(y_1,\cdots,y_n)$$

$$\times\exp\left\{\left(-\frac{1}{2\sigma^2}\right)\sum_{i=1}^n y_i^2 + \sum_{j=1}^k\frac{\beta_j}{\sigma^2}\left(\sum_{i=1}^n y_i x_{ij}\right)\right\}dy_1\cdots dy_n = 0。$$

由指数族的性质知，$Q(\beta_1,\cdots,\beta_k,\sigma^2)$ 对每个自变量的各阶偏导数都存在，而且求偏导数时可以到积分号内去做，所以有

(1.1.4)
$$\frac{\partial}{\partial\beta_u}Q = \frac{1}{\sigma^2}\int\underset{R^n}{\cdots}\int T(y_1,\cdots,y_n)\left(\sum_{i=1}^n y_i x_{iu}\right)$$

$$\times\exp\left\{\left(-\frac{1}{2\sigma^2}\right)\sum_{i=1}^n y_i^2 + \sum_{j=1}^k\frac{\beta_j}{\sigma^2}\left(\sum_{i=1}^n y_i x_{ij}\right)\right\}dy_1\cdots dy_n = 0,$$

$$u = 1,\cdots,k。$$

(1.1.5)
$$\frac{\partial^2}{\partial\beta_u\partial\beta_v}Q = \frac{1}{\sigma^4}\int\underset{R^n}{\cdots}\int T(y_1,\cdots,y_n)\left(\sum_{i=1}^n y_i x_{iu}\right)\left(\sum_{i=1}^n y_i x_{iv}\right)$$

$$\times \exp\left\{\left(-\frac{1}{2\sigma^2}\right)\sum_{i=1}^{n} y_i^2 + \sum_{j=1}^{k} \frac{\beta_j}{\sigma^2}\left(\sum_{i=1}^{n} y_i x_{ij}\right)\right\} dy_1\cdots dy_n = 0,$$

$$u,v = 1,\cdots,k。$$

(1.1.6) $$\frac{\partial}{\partial\sigma^2}Q = \frac{1}{\sigma^4}\int\cdots\int_{R^n} T(y_1,\cdots,y_n)\left(\frac{1}{2}\sum_{i=1}^{n} y_i^2 - \sum_{j=1}^{k}\beta_j\sum_{i=1}^{n} y_i x_{ij}\right)$$

$$\times \exp\left\{\left(-\frac{1}{2\sigma^2}\right)\sum_{i=1}^{n} y_i^2 + \sum_{j=1}^{k} \frac{\beta_j}{\sigma^2}\left(\sum_{i=1}^{n} y_i x_{ij}\right)\right\} dy_1\cdots dy_n = 0。$$

由(1.1.4)、(1.1.5)、(1.1.6)式可以得到：对任何 $\beta\in R^k,\sigma^2>0$ 有

$$E_{(\beta,\sigma^2)}T(y_1,\cdots,y_n)\left(\sum_{i=1}^{n} y_i^2\right)=0,$$

$$E_{(\beta,\sigma^2)}T(y_1,\cdots,y_n)\left(\sum_{i=1}^{n} y_i x_{iu}\right)\left(\sum_{i=1}^{n} y_i x_{iv}\right)=0,\quad u,v=1,\cdots,k,$$

进而得到

$$\begin{aligned}
&E_{(\beta,\sigma^2)}T(y)y'[I_n-X(X'X)^{-}X']y \\
&=E_{(\beta,\sigma^2)}T(y)y'y-E_{(\beta,\sigma^2)}T(y)y'X(a_{ij})X'y \\
&=E_{(\beta,\sigma^2)}T(y)\left(\sum_{i=1}^{n} y_i^2\right)-\sum_{u=1}^{k}\sum_{v=1}^{k}a_{uv}E_{(\beta,\sigma^2)}T(y)\left(\sum_{i=1}^{n} y_i x_{iu}\right)\left(\sum_{i=1}^{n} y_i x_{iv}\right) \\
&=0,
\end{aligned}$$

所以，

$$E_{(\beta,\sigma^2)}T(y)\hat{\sigma}^2 = 0。$$

$\hat{\sigma}^2$ 的方差有限，而且和任何一个方差有限、零的无偏估计量 $T(y)$ 的协方差是零，所以，$\hat{\sigma}^2$ 是 σ^2 的最小方差无偏估计。* □

(1.1.3)定理证明了 y 服从正态分布时，$\hat{\sigma}^2$ 是 σ^2 的最小方差无偏估计。对于一般的线性模型(1.1.1)来说，$\hat{\sigma}^2$ 还有最小方差性质吗？在这一章我们将着重研究这个问题。

本书用到的矩阵都是实矩阵，用大写字母 $A,B,\Sigma,\Psi\cdots$ 表示，m 行 n 列矩阵简称为 $m\times n$ 矩阵。矩阵的元素通常用相应的小写字母表示，例如 $A=(a_{ij})$，a_{ij}表示 A 的(i,j)元。A'表示 A 的转置。

---

* 参见[8]第二册第150页或[9]下册第51页。

$\mu(A)$表示 A 的列向量张成的向量空间。I_n 表示 n 阶单位阵，$\mu(I_n)$ 用 R^n 表示。A 是 n 阶方阵时，对角阵 $\begin{pmatrix} a_{11} & & 0 \\ & \ddots & \\ 0 & & a_{nn} \end{pmatrix}$ 用 $\text{diag}A$ 表示。$a=(a_1,\cdots,a_n)'$ 是 n 维列向量时，对角阵 $\begin{pmatrix} a_1 & & 0 \\ & \ddots & \\ 0 & & a_n \end{pmatrix}$ 用 $\text{diag}a$ 表示。$A_i, i=1,\cdots,n$ 都是 p 阶方阵时，矩阵 $\begin{pmatrix} A_1 & & 0 \\ & \ddots & \\ 0 & & A_n \end{pmatrix}$ 用 $\text{Diag}\{A_i\}$ 表示。$A\geqslant 0$ 表示 A 是对称的非负定阵。$A>0$ 表示 A 是对称的正定阵。

1.2 不变估计类

研究 $\hat{\sigma}^2$ 的最小方差性质，首先要明确 $\hat{\sigma}^2$ 和 σ^2 的哪些无偏估计量做比较，也就是要明确准备采用的估计类是什么。为了把问题简化，我们不打算广泛地采用一切 y 的二次型作为估计类，我们打算只选它的一个子集作为估计类，当然，对这个子集的选择应当是合理的。

研究 $\hat{\sigma}^2$ 的最小方差性质要涉及到 y 的二次型的方差，这需要 y 有四阶矩。为此，对线性模型(1.1.1)中的 $\varepsilon=(\varepsilon_1,\cdots,\varepsilon_n)'$ 假设

(1.2.1)　$\varepsilon_1,\cdots,\varepsilon_n$ 独立，

$$\mathrm{E}\varepsilon_i=0,\mathrm{E}\varepsilon_i^2=\sigma^2,\mathrm{E}\varepsilon_i^4=\Psi(\text{有限}),\quad i=1,\cdots,n,$$

其中 $\sigma^2>0,\Psi\geqslant\sigma^4$ 是未知参数。

首先说明，对任何指定的 $\sigma^2>0$ 及 $\Psi\geqslant\sigma^4$，满足(1.2.1)条件的随机向量 ε 都是存在的。

(1.2.2)　**例**　若 $\sigma^2>0,\Psi\geqslant\sigma^4$，则一元二次方程

$$x^2-x+\frac{\sigma^4}{\Psi+3\sigma^4}=0$$

的判别式

$$1-\frac{4\sigma^4}{\Psi+3\sigma^4}\geqslant 0,$$

方程有两个实根(或两个相同的实根),假设是 p_1 和 p_2。由于 $p_1+p_2=1$,而且 $p_1>0,p_2>0$,再取

$$a_1=-\sqrt{\frac{p_2}{p_1}}\sigma,\quad a_2=\sqrt{\frac{p_1}{p_2}}\sigma,$$

则

$$(1.2.3)\qquad \begin{pmatrix} a_1 & a_2 \\ p_1 & p_2 \end{pmatrix}$$

是一个分布列。若随机变量 $\varepsilon_1,\cdots,\varepsilon_n$ 独立,每一个 ε_i 都有分布列(1.2.3),则

$$\mathrm{E}\varepsilon_i=-\sqrt{\frac{p_2}{p_1}}\sigma p_1+\sqrt{\frac{p_1}{p_2}}\sigma p_2=0,$$

$$\mathrm{E}\varepsilon_i^2=\frac{p_2}{p_1}\sigma^2 p_1+\frac{p_1}{p_2}\sigma^2 p_2=\sigma^2,$$

$$\begin{aligned}\mathrm{E}\varepsilon_i^4&=\frac{p_2^2}{p_1^2}\sigma^4 p_1+\frac{p_1^2}{p_2^2}\sigma^4 p_2=\sigma^4\ \frac{p_1^3+p_2^3}{p_1p_2}=\sigma^4\ \frac{1-3p_1p_2}{p_1p_2}\\&=\sigma^4\left(\frac{\Psi+3\sigma^4}{\sigma^4}-3\right)=\Psi,\end{aligned}$$

$\varepsilon=(\varepsilon_1,\cdots,\varepsilon_n)'$ 就满足条件(1.2.1)。

在对(1.1.1)中的 n 维随机向量 y 增添了假设(1.2.1)后,对 y 的概率分布 P_y 虽然还不能完全了解,但是知道 P_y 属于 n 维 Borel 可测空间$(R^n,\mathscr{B}^n)$上的概率分布族

(1.2.4) $\mathscr{P}_1=\{P_\xi:\xi=X\beta+\varepsilon,\beta\in R^k,\varepsilon$ 满足条件(1.2.1)$\}$,

P_y 中有未知参数(β,σ^2,Ψ),参数空间是

(1.2.5) $\Theta_1=\{(\beta,\sigma^2,\Psi):\beta\in R^k,\sigma^2>0,\Psi\geqslant\sigma^4\}$。

若用 Λ_1 表示集合

(1.2.6) $\{(\lambda_1,\lambda_2):(\lambda_1,\lambda_2)\in R^2,\lambda_1>0,\lambda_2\geqslant\lambda_1^2\}$，

则 $\Theta_1=R^k\times\Lambda_1$。

现在考虑对未知参数的线性函数 $c'\beta+d\sigma^2+e\Psi$ 做估计的问题，(c 是已知的 k 维向量，d,e 是已知实数。)在选用估计量 $f(y)$ 估计 $c'\beta+d\sigma^2+e\Psi$ 时，考察$(R^n,\mathscr{B}^n)$上的线性变换

$$\pi_\alpha z = z + X\alpha, \quad z \in R^n,$$

这里 α 是任一指定的 k 维向量。对 n 维随机变量 y 施行线性变换 π_α，得到的

$$\pi_\alpha y = X(\beta + \alpha) + \varepsilon$$

仍是 n 维随机变量，它的概率分布 $P_{\pi_\alpha y}$仍属于 $\mathscr{P}_1$，$P_{\pi_\alpha y}$中的未知参数是$(\beta+\alpha,\sigma^2,\Psi)$，这样，$R^n$ 上的变换 π_α 导出了 Θ_1 上的一个变换 $\bar{\pi}_\alpha$

$$\bar{\pi}_\alpha(\beta,\sigma^2,\Psi) = (\beta + \alpha,\sigma^2,\Psi)。$$

所以，用 $f(y)$估计 $c'\beta+d\sigma^2+e\Psi$ 和用 $f(\pi_\alpha y)$估计 $c'(\beta+\alpha)+d\sigma^2+e\Psi$ 应当是一样的，要求

$$f(\pi_\alpha y) = f(y) + c'\alpha$$

是合理的。

变换集

(1.2.7) $\{\pi_\alpha:\pi_\alpha z=z+X\alpha,\quad \alpha\in R^k\}$

关于乘法$(\pi_{\alpha_1}\cdot\pi_{\alpha_2})z=z+X(\alpha_1+\alpha_2)$作成一个群。

(1.2.8) **定义** $f(y)$是 $c'\beta+d\sigma^2+e\Psi$ 的估计量。若对任何 $\alpha\in R^k,y\in R^n$ 有

$$f(\pi_\alpha y) = f(y) + c'\alpha,$$

则称 $f(y)$是变换群(1.2.7)之下的不变估计量。

特别，若 $f(y)$是 σ^2 的估计量，而且对任何 $\alpha\in R^k,y\in R^n$ 有

(1.2.9) $f(\pi_\alpha y)=f(y)$，

则称 $f(y)$ 是 σ^2 的关于变换群(1.2.7)的不变估计量。

记 $M=I_n-XX^+$，X^+ 表示 X 的加号逆(见附录)。容易验证，M 是对称幂等阵。

(1.2.10) **定理** (1.2.9)式对一切 $\alpha\in R^k$ 成立的充分与必要条件是 $f(y)=f(My)$。

证 条件的充分性 若 $f(y)=f(My)$，则对任何 $\alpha\in R^k$ 有

$$M\pi_\alpha y=(I-XX^+)(y+X\alpha)=(I-XX^+)y=My,$$

因而

$$f(\pi_\alpha y)=f(M\pi_\alpha y)=f(My)=f(y)。$$

条件的必要性 若对任何 $\alpha\in R^k$，(1.2.9)式成立，则对任何 y，必有

$$y=(I-XX^+)y+XX^+y=My+X(X^+y)=\pi_{X^+y}My,$$

因而

$$f(y)=f(\pi_{X^+y}My)=f(My)。\quad\square$$

(1.2.11) **推论** A 是 n 阶对称阵，σ^2 的估计量 $y'Ay$ 是变换群(1.2.7)之下的不变估计量的充要条件是 $y'Ay=yMAMy$。

前面说过，在选择 σ^2 的估计量时，要求估计量具有不变性质是合理的，我们将在 y 的二次型中选择具有不变性质的那一部分

(1.1.12) $$\mathscr{Q}_1=\{y'MAMy:A=A'\}$$

作为 σ^2 的估计类。由于

$$\hat{\sigma}^2=\frac{1}{n-r}y'[I_n-XX^+]y=y'M\left(\frac{1}{n-r}I_n\right)My\in\mathscr{Q}_1,$$

我们将在 $\mathscr{Q}_1$ 中 σ^2 的无偏估计中讨论 $\hat{\sigma}^2$ 的最小方差性质。

注 许宝騄先生在[1]中，从要求 σ^2 的二次型估计量无偏，而且方差与 β 无关导出了估计类 $\mathscr{Q}_1$。

由于

$$y'MAMy=\mathrm{tr}(MAM\cdot Myy'M),$$

My 的二次型也是 $Myy'M$ 的线性型。

$$\mathscr{Q}_1 = \{\mathrm{tr}(MAM \cdot Myy'M): A = A'\}$$

也是 $Myy'M$ 的一个线性估计类。

对于“线性估计”我们是不太陌生的，但这里的线性函数指的是随机矩阵 $Myy'M$ 元素的线性函数，这是我们不十分熟悉的。随机矩阵的线性函数也有类似于随机向量线性函数的一些性质，这些将在下一小节中看到。

1.3 随机元与射影定理

假设$(\Omega,\mathscr{A})$是一个可测空间，$(\mathscr{H},\langle\cdot,\cdot\rangle)$是一个实的，维数有限的内积空间。

(1.3.1) **定义** ξ 是 Ω 到 $\mathscr{H}$ 的一个映射，若对 $\mathscr{H}$ 中任何元 h，$\langle h,\xi\rangle$都是 $\mathscr{A}$ 可测的，则称 ξ 是$(\Omega,\mathscr{A})$上的一个随机元。

(1.3.2) **例** 取 $\mathscr{H}=R^n$，定义

$$\langle x,y\rangle = x'y, \quad \forall\ x,y \in R^n,$$

则$(\mathscr{H},\langle\cdot,\cdot\rangle)$是一个有限维实内积空间。若 $\xi=(\xi_1,\cdots,\xi_n)'$是$(\Omega,\mathscr{A})$到$(H,\langle\cdot,\cdot\rangle)$的映射，并且

(1.3.3) $\quad \sum_{i=1}^n h_i\xi_i=\langle h,\xi\rangle \quad \mathscr{A}$ 可测，$\forall\ h=(h_1,\cdots,h_n)'\in R^n$；

由(1.3.1)定义知，ξ 是$(\Omega,\mathscr{A})$上的随机元，而(1.3.3)成立的充要条件是每个 $\xi_i(i=1,\cdots,n)$都是随机变量，所以随机元 ξ 就是我们熟悉的 n 维随机向量。

(1.3.4) **例** 取

$$\mathscr{H} = \{A: A\text{ 是 } m \text{ 行 } n \text{ 列的实矩阵}\},$$

定义

$$\langle A,B\rangle = \mathrm{tr}(A'B), \quad \forall\ A,B \in \mathscr{H},$$

则$(\mathscr{H},\langle\cdot,\cdot\rangle)$是一个有限维实内积空间。若 $\xi=(\xi_{ij})$是 Ω 到 $\mathscr{H}$ 的映射，并且

(1.3.5) $\sum_{i=1}^{m}\sum_{j=1}^{n}h_{ij}\xi_{ij}=\langle h,\xi\rangle$ $\mathscr{A}$ 可测， $\forall\ h=(h_{ij})\in\mathscr{H}$，

则 ξ 是 $(\Omega,\mathscr{A})$ 上的一个随机元，而(1.3.5)成立的充要条件是每一个 $\xi_{ij}(1\leqslant i\leqslant m,1\leqslant j\leqslant n)$ 都是随机变量，所以随机元 ξ 就是我们熟知的随机矩阵。

若 ξ 是 $(\Omega,\mathscr{A})$ 到 $(\mathscr{H},\langle\cdot,\cdot\rangle)$ 的随机元，P 是 $(\Omega,\mathscr{A})$ 上的一个概率测度。对每一个 $h\in\mathscr{H}$，$\langle h,\xi\rangle$ $\mathscr{A}$ 可测，因而 $\langle h,\xi\rangle$ 是随机变量，假定 $\langle h,\xi\rangle$ 的数学期望存在(有限)，记为 $\mathrm{E}_p\langle h,\xi\rangle$。现在把 $\mathrm{E}_p\langle h,\xi\rangle$ 看作定义在 $\mathscr{H}$ 上的函数，则对任何 $h,g\in\mathscr{H}$ 及实数 c,d 有

$$\mathrm{E}_p\langle ch+dg,\xi\rangle=c\mathrm{E}_p\langle h,\xi\rangle+d\mathrm{E}_p\langle g,\xi\rangle。$$

所以，$\mathrm{E}_p\langle h,\xi\rangle$ 是 $\mathscr{H}$ 上的线性函数，有 $\mathscr{H}$ 中的唯一元 μ 使

(1.3.6) $\mathrm{E}_p\langle h,\xi\rangle=\langle h,\mu\rangle$ *

(1.3.7) **定义** 称(1.3.6)式中的 μ 为随机元 ξ 的数学期望，记为 $\mathrm{E}_p\xi$ 或 $\mathrm{E}\xi$。

(1.3.8) **例** 在(1.3.2)例中，若对任何 $a=(a_1,\cdots,a_n)'\in R^n$，$\mathrm{E}\langle a,\xi\rangle=\mathrm{E}(a'\xi)$ 存在(有限)，则 $\mathrm{E}\xi_i$ 存在(有限)，$i=1,\cdots,n$。记 $\mu=(\mathrm{E}\xi_1,\cdots,\mathrm{E}\xi_n)'$，则有

$$\mathrm{E}\langle a,\xi\rangle=\mathrm{E}\Big(\sum_{i=1}^{n}a_i\xi_i\Big)=\sum_{i=1}^{n}a_i\mathrm{E}\xi_i=\langle a,\mu\rangle,$$

由(1.3.7)定义知，$\mathrm{E}\xi=\mu=(\mathrm{E}\xi_1,\cdots,\mathrm{E}\xi_n)'$，可见，随机元 ξ，也就是 n 维随机变量 ξ 的数学期望的定义与通常的定义是一致的。

(1.3.9) **例** 在(1.3.4)例中，若对任何 $A=(a_{ij})\in\mathscr{H}$，$\mathrm{E}\langle A,\xi\rangle=\mathrm{Etr}(A'\xi)$ 存在(有限)，则 $\mathrm{E}\xi_{ij}$ 存在(有限)，$i=1,\cdots,m;j=1,\cdots,n$。记 $\mu=(\mathrm{E}\xi_{ij})$ 则有

$$\mathrm{E}\langle A,\xi\rangle=\mathrm{E}\sum_{i=1}^{m}\sum_{j=1}^{n}a_{ij}\xi_{ij}=\sum_{i=1}^{m}\sum_{j=1}^{n}a_{ij}\mathrm{E}\xi_{ij}=\langle A,\mu\rangle,$$

* 见[10]第164页。

可见，随机元 ξ——m 行 n 列随机矩阵 ξ 的数学期望的定义与通常的定义是一致的。

下面我们定义随机元的协方差。

若 ξ 是 $(\Omega,\mathscr{A},P)$ 到 $(\mathscr{H},\langle\cdot,\cdot\rangle)$ 的随机元，而且对任何 g，$h\in\mathscr{H}$，随机变量 $\langle g,\xi\rangle$ 和 $\langle h,\xi\rangle$ 的协方差 $\mathrm{cov}_p(\langle g,\xi\rangle,\langle h,\xi\rangle)$ 存在有限，则

(1) 对任何 $g_1,g_2,h\in\mathscr{H}$ 及实数 c_1,c_2 有

$$\begin{aligned}&\mathrm{cov}_p(\langle c_1g_1+c_2g_2,\xi\rangle,\langle h,\xi\rangle)\\&=c_1\,\mathrm{cov}_p(\langle g_1,\xi\rangle,\langle h,\xi\rangle)+c_2\,\mathrm{cov}_p(\langle g_2,\xi\rangle,\langle h,\xi\rangle),\end{aligned}$$

(2) 对任何 $g,h_1,h_2\in\mathscr{H}$ 及实数 d_1,d_2 有

$$\begin{aligned}&\mathrm{cov}_p(\langle g,\xi\rangle,\langle d_1h_1+d_2h_2,\xi\rangle)\\&=d_1\,\mathrm{cov}_p(\langle g,\xi\rangle,\langle h_1,\xi\rangle)+d_2\,\mathrm{cov}_p(\langle g,\xi\rangle,\langle h_2,\xi\rangle)。\end{aligned}$$

所以，$\mathrm{cov}_p(\langle g,\xi\rangle,\langle h,\xi\rangle)$ 是 $\mathscr{H}$ 上的双线性函数，有唯一确定的 $\mathscr{H}$ 到 $\mathscr{H}$ 的线性变换 Σ，使得

(1.3.10) $\mathrm{cov}_p(\langle g,\xi\rangle,\langle h,\xi\rangle)=\langle g,\Sigma h\rangle$*， $\forall\, g,h\in\mathscr{H}$。

由协方差的性质知，Σ 是非负定的。

(1.3.11) **定义** 称(1.3.10)式中的 Σ 为随机元 ξ 的协方差算子，记为 $\mathrm{Cov}_p\xi$ 或 $\mathrm{Cov}\xi$。

(1.3.12) **例** 在(1.3.2)例中，若对任何 $g,h\in R^n$，$\mathrm{cov}(g'\xi,h'\xi)$ 存在有限，则 $\mathrm{cov}(\xi_i,\xi_j)$ 存在有限，$i,j=1,\cdots,n$。记

$D=(\mathrm{cov}(\xi_i,\xi_j))$，则有

$$\mathrm{cov}(g'\xi,h'\xi)=g'Dh,\quad \forall\, g,h\in R^n,$$

由(1.3.11)定义知 ξ 的协方差算子是 D，与 ξ 的协方差阵一致。

若 ξ 是 $(\Omega,\mathscr{A})$ 到 $(\mathscr{H},\langle\cdot,\cdot\rangle)$ 的随机元，$\{P_\theta:\theta\in\Theta\}$ 是 $(\Omega,\mathscr{A})$ 上的一族概率测度，Θ 是参数空间，对每一个 $\theta\in\Theta$，$\mathrm{E}_{P_\theta}\xi$ 及 $\mathrm{Cov}_{P_\theta}\xi$ 存在有限，简记为 $\mathrm{E}_\theta\xi$ 及 $\mathrm{Cov}_\theta\xi$。

---

* 见[10]第 241 页。