国外名校最新教材精选

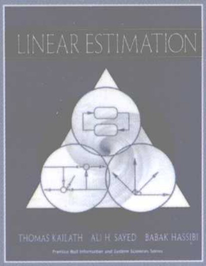
Thomas Kailath　Ali H. Sayed　Babak Hassibi

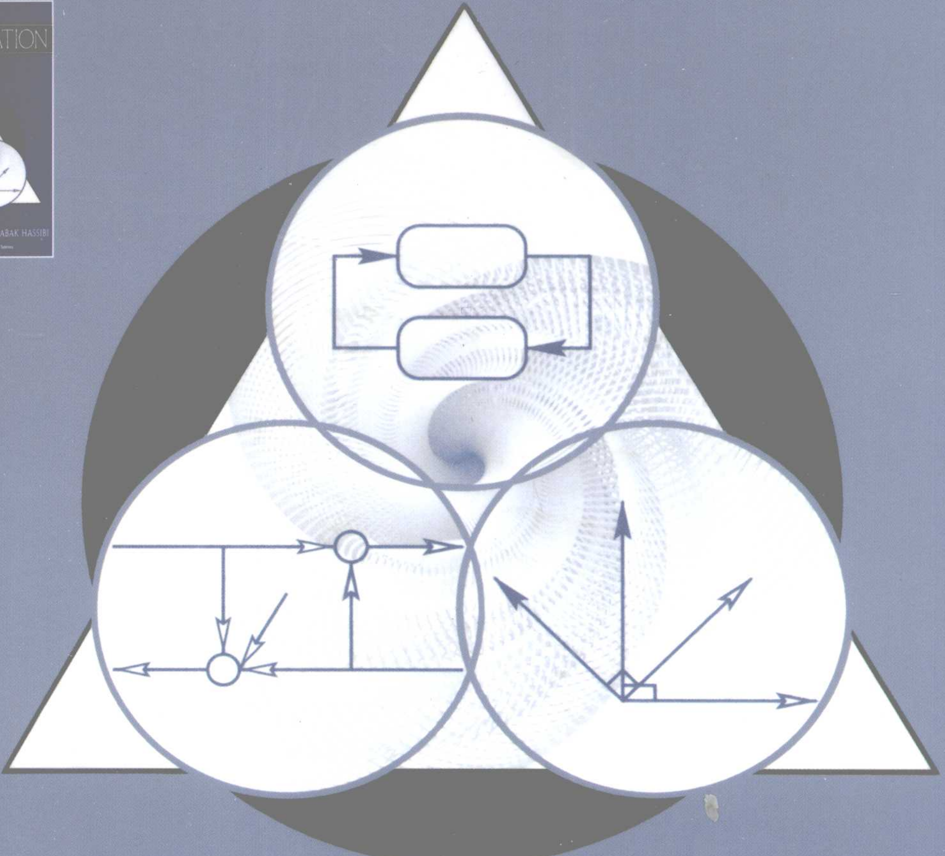
线性估计

Linear Estimation

托马斯・凯拉斯
〔美〕阿里・H・赛义德　著
巴巴克・哈斯比

（影印版）





西安交通大学出版社
XI'AN JIAOTONG UNIVERSITY PRESS

图书在版编目(CIP)数据

线性估计=Linear Estimation:英文 / (美)凯拉斯(Kailath,T.), (美)赛义德(Sayed,A. H.),(美)哈斯比(Hassibi,B.)著. —影印本. —西安:西安交通大学出版社,2008.11

(国外名校最新教材精选)

ISBN 978-7-5605-2949-3

Ⅰ.线…　Ⅱ.①凯…②赛…③哈…　Ⅲ.线性估计-高等学校-教材-英文　Ⅳ.O231.1

中国版本图书馆 CIP 数据核字(2008)第 170792 号

书　　名	线性估计(影印版)
著　　者	〔美〕托马斯·凯拉斯　阿里·H·赛义德　巴巴克·哈斯比
策划编辑	赵丽平
责任编辑	李　颖

出版发行	西安交通大学出版社 (西安市兴庆南路 10 号　邮政编码 710049)
网　　址	http://www.xjtupress.com
电　　话	(029)82668357　82667874(发行中心) (029)82668315　82669096(总编办)
传　　真	(029)82668280
印　　刷	西安交通大学印刷厂

开　　本	787mm×1092mm　1/16	印　　张	55
印　　数	0001～2000	字　　数	1016 千字
版次印次	2008 年 12 月第 1 版		2008 年 12 月第 1 次印刷
书　　号	ISBN 978-7-5605-2949-3/O·285		
定　　价	98.00 元		

读者购书、书店添货如发现印装质量问题，请与本社发行中心联系、调换。

订购热线：(029)82665248　(029)82665249

投稿热线：(029)82665380

读者信箱：banquan1809@126.com

推荐者序

呈现在读者面前的《线性估计》(Linear Estimation)这本书,是世界著名学者Kailath(凯拉斯)等编著,由Prentice Hall出版的信息与系统科学丛书之一。

信息与系统科学的精髓,在于处理各种不确定的数据或信息,从而达到对各种系统运行状态估计的目的。因而,研究不确定性理论,特别是估计理论具有非常重要的科学意义。该书囊括了线性估计理论发展至今所有的前沿知识,是一本体系完整、论述严谨、引领信息与系统科学未来发展的一本名著。全书由17章正文和7章附录构成。

第1章是概论部分,讲述渐近观测器和最优瞬态观测器,以及未来具有吸引力的研究方向,包括平滑估计、时变模型的扩展、时不变系统的快速算法、数值方法、阵列计算等;还介绍了新息过程和系统的稳态行为,最后还介绍了相关的几个问题,包括自适应递推最小二乘滤波、线性二次控制、H_∞估计、H_∞自适应滤波、H_∞控制,以及线性代数与矩阵论等。

第2章讲述确定性最小二乘问题。与传统的著作不同,该书非常深入地讨论了最小二乘判据、最小二乘的经典解、几何描述的正交性条件、正则化的最小二乘问题、阵列求解算法、递推最小二乘算法、总体最小二乘算法等。第3章讲述随机最小二乘问题。详细描述了随机估计问题的本质;给出了线性最小均方估计器的描述和分析方法;讨论了几何描述的正交条件;而且讨论了线性模型,包括特定条件下的信息形式、Gauss-Markov定理、联合估计器、与确定性最小二乘的等价等。

第4章讨论新息过程。先从随机过程的估计问题入手,讨论固定区间平滑问题、因果性滤波问题、Wiener-Hopf方程等,再用几何方法和代数方法分别描述新息过程,讨论了修正的Gram-Schmidt直交化方法、给定新息过程的估计问题,以及利用新息方法的滤波问题等;深入研究了确定性最小二乘问题的新息方法,以及指数相关过程等。第5章研究状态空间模型,进一步讨论指数相关过程中的有限区间问题、平稳性的初始条件、过程模型的信息等,还进一步讨论了非平稳过程。与此同时,深入研究了高阶过程与状态空间模型的关系,包括自回归过程、初始条件、状态空间描述、标准状态空间模型、前向Markov模型和后向Markov模型,以及Markov表现与标准状态空间模型的关系等。第6章研究平稳过程的新息,包括通过谱分解得到的新息、信号与系统的相关知识、平稳随机过程的特性、典范谱分解、标量z谱、向量值平稳过程,以及连续时间系统与过程等。

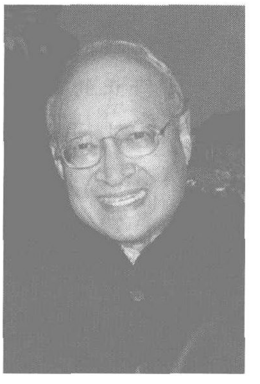
第7章讨论标量过程的Wiener理论,包括连续时间的Wiener平滑、连续时间的Wiener-Hopf方程、对应的离散时间的Wiener平滑和Wiener-Hopf方程等,同时探讨了向量过程的Wiener理论。第8章研究递推Wiener滤波问题,包括时不变状态空间模型分析、典范谱分解、给定状态空间模型的递推估计、递推Wiener滤波器、时变模型的扩展等。第9章研究Kalman滤波问题,包括标准状态空间模型、获取新息的

Kalman 滤波递推算法、用于预报和滤波的状态估计器递推算法、协方差阵及其逆的三角分解、重要的谱假设、基于协方差的滤波器、近似非线性滤波、后向 Kalman 滤波递推算法等。第 10 章讨论平滑估计器问题,包括一般的平滑公式、状态空间结构、Rauch-Tung-Striebel(RTS)递推算法、双滤波器公式、Hamilton 方程及其变种等。

第 11 章讨论快速算法,包括零初始条件和平稳过程的情况,以及推广到时变系统等。第 12 章专门研究阵列算法,包括对标量量测更新的 Potter 算法,以及几种典型的阵列算法,后者包括标准假设、时间更新、量测更新、预测估计、滤波估计、估计更新等;同时给出了算法实例和推导证明;进而讨论阵列算法的 Paige 形式、信息形式的阵列算法、平滑的阵列算法等。第 13 章专门讨论快速阵列算法,由形成阵列算法的显式方程开始,推广到时变系统等。第 14 章分析算法的渐近行为,即讨论各种算法的收敛性以及解的特征;同时也分析了指数收敛与快速递推的问题。

第 15 章是理论研究,分析估计与控制问题的对偶性和等价性。其中包括对偶基、线性模型的应用、对偶性与等价关系、因果约束下的对偶性量测约束与分离原理、频域的对偶性、补状态空间模型等。第 16 章研究连续时间状态空间的估计问题,包括连续时间模型、给定状态空间和协方差阵时的 Kalman 滤波方程、利用新息过程的直接求解、平滑估计器、时不变模型的快速算法、算法的渐近行为、稳态滤波器等。第 17 章讨论扩散理论法,包括广义发射线模型、反向演化、各种 Riccati 公式、时不变模型与均匀媒体、离散时间的扩散公式等。

附录 A 是一些有用的矩阵论结果;附录 B 是酉变换与 J-酉变换;附录 C 介绍某些系统理论的概念;附录 D 介绍 Lyapunov 方程;附录 E 介绍代数 Riccati 方程;最后附录 F 介绍置换结构。



该书作者 Thomas Kailath 教授 1935 年 6 月出生在印度西部城市浦那,1957 年移居美国,分别于 1959 年和 1961 年在 MIT 获得电气工程的硕士和博士学位。1962 年以后在 MIT 的通信系统分部工作,1963 年起在 Stanford 大学任教,1968 年晋升教授,1970 年起为 IEEE Fellow;1971 年—1981 年任 Stanford 大学信息系统实验室主任,是世界著名的控制与系统科学专家,是美国科学院和工程院院士、第三世界科学院院士和印度工程院院士。他的研究兴趣涉及信息理论、通信系统、计算、控制、线性系统、统计信号处理、大规模集成电路等。他的名著《线性系统理论》(Linear System Theory, Springer-Verlag, 1991)风靡全球,在我国自动控制界有重要影响,对推动我国控制科学与工程学科的发展产生了重要作用。

我们深信,本书的影印出版,对于我国信息科学领域的学科发展必将产生更大的推动作用。

韩崇昭

2008 年 11 月于西安交通大学

To our parents and families

Preface

The problem of estimating the values of a random (or stochastic) process given observations of a related random process is encountered in many areas of science and engineering, *e.g.*, communications, control, signal processing, geophysics, econometrics, and statistics. Although the topic has a rich history, and its formative stages can be attributed to illustrious investigators such as Laplace, Gauss, Legendre, and others, the current high interest in such problems began with the work of H. Wold, A. N. Kolmogorov, and N. Wiener in the late 1930s and early 1940s. N. Wiener in particular stressed the importance of modeling not just "noise" but also "signals" as random processes. His thought-provoking originally classified 1942 report, released for open publication in 1949 and now available in paperback form under the title *Time Series Analysis*, is still very worthwhile background reading.

As with all deep subjects, the extensions of these results have been very far-reaching as well. A particularly important development arose from the incorporation into the theory of multichannel state-space models. Though there were various earlier partial intimations and explorations, especially in the work of R. L. Stratonovich in the former Soviet Union, the chief credit for the explosion of activity in this direction goes to R. E. Kalman, who also made important related contributions to linear systems, optimal control, passive systems, stability theory, and network synthesis.

In fact, least-squares estimation is one of those happy subjects that is interesting not only in the richness and scope of its results, but also because of its mutually beneficial connections with a host of other (often apparently very different) subjects. Thus, beyond those already named, we may mention connections with radiative transfer and scattering theory, linear algebra, matrix and operator theory, orthogonal polynomials, moment problems, inverse scattering problems, interpolation theory, decoding of Reed–Solomon and BCH codes, polynomial factorization and root distribution problems, digital filtering, spectral analysis, signal detection, martingale theory, the so-called $\mathcal{H}_\infty$ theories of estimation and control, least-squares and adaptive filtering problems, and many others. We can surely apply to it the lines written by William Shakespeare about another (beautiful) subject:

> "Age does not wither her, nor custom stale,
> Her infinite variety."

Though we were originally tempted to cover a wider range, many reasons have led us to focus this volume largely on estimation problems for finite-dimensional linear systems with state-space models, covering most aspects of an area now generally known as Wiener and Kalman filtering theory. Three distinctive features of our treatment are the pervasive use of a geometric point of view, the emphasis on the numerically favored square-root/array forms of many algorithms, and the emphasis on equivalence and duality concepts for the solution of several related problems in adaptive filtering, estimation, and control. These features are generally absent in most prior treatments, ostensibly on the grounds that they are too abstract and complicated. It is our hope that these misconceptions will be dispelled by the presentation herein, and that the fundamental simplicity and power of these ideas will be more widely recognized and exploited.

The material presented in this book can be broadly categorized into the following topics:

- Introduction and Foundations
  - Chapter 1: Overview
  - Chapter 2: Deterministic Least-Squares Problems
  - Chapter 3: Stochastic Least-Squares Problems
  - Chapter 4: The Innovations Process
  - Chapter 5: State-Space Models
- Estimation of Stationary Processes
  - Chapter 6: Innovations for Stationary Processes
  - Chapter 7: Wiener Theory for Scalar Processes
  - Chapter 8: Recursive Wiener Filters
- Estimation of Nonstationary Processes
  - Chapter 9: The Kalman Filter
  - Chapter 10: Smoothed Estimators
- Fast and Array Algorithms
  - Chapter 11: Fast Algorithms
  - Chapter 12: Array Algorithms
  - Chapter 13: Fast Array Algorithms
- Continuous-Time Estimation
  - Chapter 16: Continuous-Time State-Space Estimation
- Advanced Topics
  - Chapter 14: Asymptotic Behavior
  - Chapter 15: Duality and Equivalence in Estimation and Control
  - Chapter 17: A Scattering Theory Approach

Being intended for a graduate-level course, the book assumes familiarity with basic concepts from matrix theory, linear algebra, linear system theory, and random processes. Four appendices at the end of the book provide the reader with background material in all these areas.

There is ample material in this book for the instructor to fashion a course to his or her needs and tastes. The authors have used portions of this book as the basis for one-quarter first-year graduate level courses at Stanford University, the University of California at Los Angeles, and the University of California at Santa Barbara; the students were expected to have had some exposure to discrete-time and state-space theory. A typical course would start with Secs. 1.1–1.2 as an overview (perhaps omitting the matrix derivations), with the rest of Ch. 1 left for a quick reading (and re-reading from time to time), most of Chs. 2 and 3 (focusing on the geometric approach) on the basic deterministic and stochastic least-squares problems, Ch. 4 on the innovations process, Secs. 6.4–6.5 and 7.3–7.7 on scalar Wiener filtering, Secs. 9.1–9.3, 9.5, and 9.7 on Kalman filtering, Secs. 10.1–10.2 as an introduction to smoothing, Secs. 12.1–12.5 and 13.1–13.4 on array algorithms, and Secs. 16.1–16.4 and 16.6 on continuous-time problems.

More advanced students and researchers would pursue selections of material from Sec. 2.8, Chs. 8, 11, 14, 15, and 17, and Apps. E and F. These cover, among other topics, least-squares problems with uncertain data, the problem of canonical spectral factorization, convergence of the Kalman filter, the algebraic Riccati equation, duality, backwards-time and complementary models, scattering, etc. Those wishing to go on to the more recent $\mathcal{H}_\infty$ theory can find a treatment closely related to the philosophy of the current book (*cf.* Sec. 1.6) in the research monograph of Hassibi, Sayed, and Kailath (1999).

A feature of the book is a collection of nearly 300 problems, several of which complement the text and present additional results and insights. However, there is little discussion of real applications or of the error and sensitivity analyses required for them. The main issue in applications is constructing an appropriate model, or actually a set of models, which are further analyzed and then refined by using the results and algorithms presented in this book. Developing good models and analyzing them effectively requires not only a good appreciation of the actual application, but also a good understanding of the theory, at both an analytical and intuitive level. It is the latter that we have tried to achieve here; examples of successful applications have to be sought in the literature, and some references are provided to this end.

Acknowledgments

The development of this textbook has spanned many years. So the material, as well as its presentation, has benefited greatly from the inputs of the many bright students who have worked with us on these topics: J. Omura, P. Frost, T. Duncan, R. Geesey, D. Duttweiler, H. Aasnaes, M. Gevers, H. Weinert, A. Segall, M. Morf, B. Dickinson, G. Sidhu, B. Friedlander, A. Vieira, S. Y. Kung, B. Levy, G. Verghese, D. Lee, J. Delosme, B. Porat, H. Lev-Ari, J. Cioffi, A. Bruckstein, T. Citron, Y. Bresler, R. Roy, J. Chun, D. Slock, D. Pal, G. Xu, R. Ackner, Y. Cho, P. Park, T. Boros, A. Erdogan, U. Forsell, B. Halder, H. Hindi, V. Nascimento, T. Pare, R. Merched, and our young friend Amir Ghazanfarian (*in memoriam*) from whom we had so much more to learn.

We are of course also deeply indebted to the many researchers and authors in this beautiful field. Partial acknowledgment is evident through the citations and references; while the list of the latter is quite long, we apologize for omissions and inadequacies arising from the limitations of our knowledge and our energy. Nevertheless, we would be remiss not to explicitly mention the inspiration and pleasure we have gained in studying the papers and books of N. Wiener, R. E. Kalman, and P. Whittle.

Major support for the many years of research that led to this book was provided by the Mathematics Divisions of the Air Force Office of Scientific Research and the Army Research Office, by the Joint Services Electronics Program, by the Defense Advanced Research Projects Agency, and by the National Science Foundation. Finally, we would like to thank Bernard Goodwin and Tom Robbins, as well as the staff of Prentice Hall, for their patience and other contributions to this project.

T. Kailath
Stanford, CA

A. H. Sayed
Westwood, CA

B. Hassibi
Murray Hill, NJ

Symbols

We collect here, for ease of reference, a list of the main symbols and signs used throughout the text.

$\mathbb{R}$	The set of real numbers.
$\mathbb{C}$	The set of complex numbers.
$.^T$	Matrix transposition.
$.^*$	Complex conjugation; Hermitian transposition.
$\diamond$	denotes the end of a theorem/lemma/proof/example/remark.
z^{-1}	denotes a unit-time delay.
$\alpha \in \mathcal{S}$	The element α belongs to the set $\mathcal{S}$.
$\mathbf{x}$	a boldface letter denotes a random variable.
x	a letter in normal font denotes a vector in Euclidean space.
$E\mathbf{x}$	denotes the expected value of a random variable $\mathbf{x}$.
$\langle \mathbf{x}, \mathbf{y} \rangle$	denotes $E\mathbf{x}\mathbf{y}^*$ for column random vectors $\mathbf{x}$ and $\mathbf{y}$.
$\|\mathbf{x}\|^2$	denotes $E\mathbf{x}\mathbf{x}^*$ for a zero-mean random variable $\mathbf{x}$.
$\mathbf{x} \perp \mathbf{y}$	denotes uncorrelated zero-mean random variables $\mathbf{x}$ and $\mathbf{y}$.

$\langle x, y\rangle$	denotes the inner product x^*y for column vectors x and y.
$\|x\|^2$	denotes x^*x for a column vector x.
$\|x\|$	denotes $\sqrt{x^*x}$ for a column vector x.
$x \perp y$	denotes orthogonal vectors x and y.
$\|A\|_2$	The 2-induced norm = the maximum singular value of A.
$\|A\|_{\rm F}$	The Frobenius norm of A.
$a \triangleq b$	The quantity a is defined as b.
$a \propto b$	The quantity a is proportional to b.
$\text{col}\{a, b\}$	a column vector with entries a and b.
$\text{vec}\{A\}$	a column vector formed by stacking the columns of A.
$\text{diag}\{a, b\}$	a diagonal matrix with diagonal entries a and b.
$a \oplus b$	The same as $\text{diag}\{a, b\}$.
0	a zero scalar, vector, or matrix.
I_n	The identify matrix of size $n \times n$.
$x(z)$ or $X(z) = \mathcal{Z}\{x_i\}$	denotes the bilateral z-transform of a sequence $\{x_i\}$.
$X(f) = \mathcal{F}\{x(t)\}$	denotes the Fourier transform of a function $x(t)$.
$X(s) = \mathcal{L}\{x(t)\}$	denotes the bilateral Laplace transform of $x(t)$.
$X(e^{j\omega})$	denotes the Discrete-Time Fourier Transform of $\{x_i\}$.
$\mathcal{L}\{\mathbf{x}_1, \mathbf{x}_2, ..\}$	denotes the linear span of the variables $\{\mathbf{x}_1, \mathbf{x}_2, \ldots\}$.
$\hat{\mathbf{x}}_{i\|j}$	l.l.m.s. estimator of $\mathbf{x}_i$ given observations up to time j.
$\hat{\mathbf{x}}_i$	l.l.m.s. estimator of $\mathbf{x}_i$ given observations up to time $i-1$.
$\tilde{\mathbf{x}}_{i\|j}$	The estimation error $\mathbf{x}_i - \hat{\mathbf{x}}_{i\|j}$.
$\tilde{\mathbf{x}}_i$	The estimation error $\mathbf{x}_i - \hat{\mathbf{x}}_i$.
$\{\cdot\}_+$	Causal part of a transfer function.
$\{\cdot\}_-$	Anti-causal part of a transfer function.
$\{\cdot\}_{s.c.}$	Strictly causal part of a transfer function.

$P > 0$	a positive-definite (p.d.) matrix P.
$P \geq 0$	a positive-semidefinite (p.s.d.) matrix P.
$P^{1/2}$	a square-root factor of a matrix $P \geq 0$, usually triangular.
$A > B$	means that $A - B$ is positive-definite.
$A \geq B$	means that $A - B$ is positive-semidefinite.
$\det A$	Determinant of the matrix A.
trace A	Trace of the matrix A.
$O(n)$	A constant multiple of n, or of the order of n.
QR	The QR factorization of a matrix.
LDU	Lower-diagonal-upper decomposition of a matrix.
UDL	Upper-diagonal-lower decomposition of a matrix.
LDL^*	LDU decomposition of a Hermitian matrix.
UDU^*	UDL decomposition of a Hermitian matrix.
Thm.	"Theorem."
Cor.	"Corollary."
Def.	"Definition."
Fig.	"Figure."
LTI	"Linear time-invariant."
l.l.s.	"linear least-squares."
l.l.m.s.	"linear least-mean-squares."
l.l.m.s.e	"linear least-mean-squares estimation/estimator."
m.m.s.e.	"minimum mean-square error."
LS	"least-squares."
p.d.f.	"probability density function."
iff	"if and only if."
a.e.	"almost everywhere."
w.r.t.	"with respect to."
RHS	"right-hand side."
LHS	"left-hand side."

ROC	"Region of convergence."
ARE	"Algebraic Riccati equation."
DARE	"Discrete-time algebraic Riccati equation."
CARE	"Continuous-time algebraic Riccati equation."
LMI	"Linear Matrix Inequality."
AR	"Autoregressive model."
MA	"Moving average model."
ARMA	"Autoregressive moving average model."
FIR	"Finite impulse response filter."
IIR	"Infinite impulse response filter."
SNR	"Signal to noise ratio."
MAP	"Maximum a-posteriori."
EKF	"Extended Kalman filter."
SISO	"Single-input single-output."
MIMO	"Multiple-input multiple-output."

Contents