应用最优估计

A 盖尔布主编 蔡世忠 张最良等译

中国人民解放军国防科学技术大学

应用最优估計

A. 盖尔布 主编

蔡世忠 张最良 张祁康
吴成璠 薛沛丰 译

张最良 沙钰 校

中国人民解放军国防科学技术大学

应用最优估计

A. 盖尔布 主编

蔡世忠 张最良 等译

*

编 号：303010181 印 数：1—1300 册

出 版：国防科学技术大学 印 刷：本校印刷厂

1980年7月出版 成本费：2.86元

内 容 简 介

《应用最优估计》是一本介绍现代估计理论及其应用的比较优秀的著作。

全书由三部分组成：一、介绍有关的基本数学知识；二、介绍最优线性滤波、平滑和非线性估计的基本理论；三、讨论最优估计算法的设计、亚优滤波器的设计、灵敏度分析、计算机实现等问题。此外，还介绍了估计方法的改进、最优线性估计理论与观测器理论及最优控制理论的联系等问题。

本书理论结合实践，着重阐述物理实质和实际应用，并列举了一百多个例题和习题。

本书适于自动控制、惯性导航、无线电通讯、计算数学、系统工程等领域的科技人员参考。亦可供高等院校有关专业高年级生、研究生作为教学用书。

译 者 前 言

《应用最优估计》是一本介绍现代估计理论及其应用的比较优秀的著作。作者成功地把理论结合于实际，总结了许多应用最优估计理论的实践经验，并且写得深入浅出、着重于阐述物理实质和实际应用中的重要问题。书中还列举了內容丰富的例题习题，其中不少题目本身就是从实际研究中提炼出来的，便于读者自学和具体应用。

目前，我国进入了社会主义现代化建设的新时期，有关科技人员正在学习和应用现代估计理论。译者认为本书将有助于读者“达到有能力参与实际估计器的设计和计算的水平”，是一本比较成熟、比较适用的近著。在校有关部门的支持下，完成了译校和出版工作，使本书能尽快地和读者见面。

本书译者：第一章张礽康、第二、六章张最良，第三、七章吳成璠、第四、五章薛沛丰，第八、九章蔡世忠；校阅：张最良。

在此，感谢沙钰同志仔细地阅读了译稿，并提出了有益的意见、侯云同志为本书设计了封面。此外，还有不少同志为本书的出版作了许多工作，谨表示我们的谢意。由于我们水平有限，译校中错误不当之处，欢迎批评指正。

前　言

从数据中提取信息的过程叫做估计。这些数据是可以用来推断信息的，并可能含有误差。现代估计方法，考虑到测量误差，干扰和控制作用对系统的影响，以及信息的验前知识，利用已知的关系从测量值计算所要求的信息。不同的测量可以混合起来形成“最好”的估计，而测不到的信息可以以最优方式近似得到。本书的意图是使读者能够达到有能力参与实际估计器的设计和计算的水平。所以，内容着重于最优估计的应用方面而不是理论方面。自始至终我们的意图是对现代估计理论和实践的基本问题给以简单而有趣的描绘。我们不去追求理论上的完美性而着重启发式的论证，特别着重于物理上的理解以及对实用很重要的关键问题。

全书由三个主要部份组成。第一部分引入主题并提供有关基本数学知识的简短介绍。其中第一章是一个包括历史回顾的简短引言；第二、第三章是有关随机过程理论和线性动态系统状态空间特性的基本数学知识，这两者都是了解最优估计理论的重要前提。第二部分提供与最优估计理论有关的推导、解释和举例。第四、第五章分别介绍最优线性滤波和平滑问题，而第六章介绍非线性滤波和平滑问题。第三部分讨论若干实际问题，它们往往关系到所实现的最优估计器的成败。第七、第八章以适当篇幅讨论亚优滤波、灵敏度分析和实现考虑等实用中非常重要的问题。第九章提供其它有实际意义的课题，包括估计理论的改进和其它见解、最优线性估计理论与最优线性控制理论在数学上的密切关系等。

贯穿全书列举了许多例题，有助于更好地说明理论内容。另外，每章结尾附加“插入”答案的习题，供读者进一步自学。

本书是分析科学公司(The Analytic Sciences Corporation——缩写为TASC)在若干美国政府机关讲课的成果。讲课的原稿又是以TASC把现代估计理论应用于各类大规模系统所得到的大量实际经验为基础形成的。因此，实际上TASC技术部所有成员都在这一时期或那一时期对本书所包含的材料做出了贡献。这里，特别要感谢TASC中除主要作者外直接对编写本书做出贡献的成员。Bard S. Crawford和William F O'Halloran, Jr. 提供了整节的内容；Robert G. Bellaire, Julian L. Center, Jr., Joseph A. D'Appolito, Norman H. Josephy（现在威斯康辛大学）；Bahar J. Uttam和Ronald S. Warren提供了正文补充、技术注释和各种不同性质的理解。Frances A. Smith打印了大多数原稿，William R. Sullivan和Vicky M. koczerga完成了所有的艺术工作。Willie J. Smith为澄清错误通读了原稿并组织了本书的付印工作。麻省理工学院（MIT）的教授Renwick E. Curry, John J. Deyst, Jr.和Wallace E.Vander Velde通过技术讨论和提供包括在若干章末尾的一些习题为本书作出了贡献。

马萨诸塞大学的教授 charles E. Hutchinson 由于早期参与 TASC 的工作也为本书做出了贡献。最后，特别要感谢 TASC 的 Harry B. Silverman 从本书计划一开始就给予热情的鼓励。

Arthur Gelb

1974.1.14.

目　　录

第一章 引 言

历 史 回 顾

与随机变量有关的数据处理方法的产生可以追溯到高斯(Gauss)（约1800年），他发明了确定性最小平方技术，并把它用于一个比较简单的轨道测量问题（文献1）。对估计理论的广泛问题做出的下一个重要贡献出现在100年后，当时，从事概率密度函数研究的费歇(Fisher)（约1910年)引入了极大似然估计方法（文献2）。利用随机过程理论，維纳(Wiener)（约1940年)提出了一个在频率域设计统计最优滤波器的方法(文献3,4)*此技术适用于以相关函数和連续滤波器脉冲响应表示的連续时间问题。它局限于统计平稳过程，并只提供稳态的最优估值。在同一时期内柯尔莫柯洛夫(Kolmogorov)处理了离散时间问题（文献5）。在以后的廿年中，維纳的工作以一种经常要求烦复计算的方法被扩展到包括非平稳和多維的系统（文献6—8）。卡尔曼(Kalman)和其他一些人(1960年前后）基于状态空间时间域的公式，提出了最优递推滤波技术(文献9—13)。这种方法就是现在著名的卡尔曼滤波器，它理想地适于用数字计算机实现。确实，在现代多敏感器系统中，这正是数据混合（处理）的基础。

看到高斯的工作与“现代”的方法之间有许多相似之处是很有意思的。如文献14所指出的，高斯注意到用冗余的数据来消除测量误差影响的必要性；他提出了在研究中建立系统动力学模型的问题；他提到观测的不精确性，并因此而开始了概率方法的阶段；他提到观测的“适宜组合”将提供最精确的估计，由此接触到了估计器结构和性能判据定义的问题；他提到为了确定未知量而必须要求的观测的数目，这样就提出了当前称之为系统“可观测性”的问题。还可以引证其他类似之处。事实上可以证明，卡尔曼滤波器本质上就是对高斯的原始最小平方问题的一个递推解*。

同样令人感兴趣的是在古典技术和现代技术之间的两个基本区别。即，用随机过程代替确定性的信号描述，以及用高速数字计算机产生数值解代替用“笔和纸”来求解而要求的闭合形式解。前一个考虑使得现代数学能精确地描述要研究的物理现象，而后者极大地扩充了可以研究的问题的范围。

最 优 估 计

一个最优估计器是一个计算算法，它利用：关于系统和测量动力学的知识，假设的系统噪声和测量误差的统计特性，以及初始条件信息，对测量值进行处理以得到关于系统状态的最小误差估计。**这类数据处理器的予想的优点包括它在明确定义的统计意义

* 递推解指能夠顺序地而不是成批地处理测量数据。

** 依照某种提法的最优性判据。

上使估值误差最小，以及它利用了所有的测量数据和有关系统的验前知识。相应的潜在缺点是它对错误的验前模型和统计特性的灵敏性，以及特有的计算复杂性。在这些叙述中所包含的重要概念将在下面详细探讨。

图1.0-1描写了三种我们感兴趣的估计问题。当我们希望得到其状态估值的时间与最后测量时间重合时，称此问题为滤波问题；当希望得到的状态估值的时间处于可得到测量数据的时间间隔之内时，就作为平滑问题来处理；而当感兴趣的状态估值的时间在最后得到的测量时间之后时，就称为予测问题。

或许最普通的最优滤波技术是由卡尔曼提出(文献10)的估计线性系统状态的技术。此项技术，如图 1.0-2 所示，提供了一个适宜的例子，可以说明最优估计器的能力和限制。例如，给定一个线性系统模型和对它的特性的任意测量，加上表征系统和测量误差的统计模型，以及初始条件信息，卡尔曼滤波器描述如何去处理测量数据。然而，卡尔曼滤波器本身并不解决建立最优测量程序的问题，它也不解决当存在参数不确定性时的设计问题，以及如何处理计算误差的问题。除了用于推导滤波算法的设计判据外，为了解决问题还必须加上其他的设计判据。

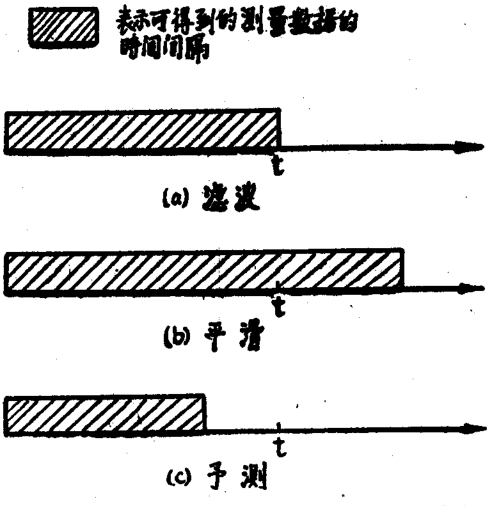

图 1.0-1 估计问题的三种类型（要求在时间 t 估计）

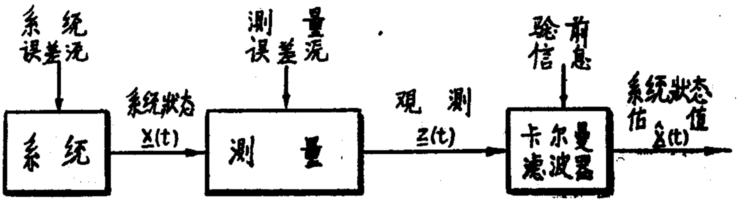

图 1.0-2 描述系统、测量和估计器的方块图

最优估计理论的应用

最优估计理论具有非常广泛的应用范围。例如：核医学的示踪研究，统计图象处理，交通密度估计，化学过程控制，河流水量估计，发电站系统负载予测，向量心电图的分类，卫星轨道估计以及核反应堆参数识别。

估计问题可以就单个敏感器测量单个过程来提出，也可以就多个敏感器测量多个过程来提出。后一种情况称为**多敏感器系统**。假设有 l 个敏感器，它们对 m 个物理过程提供测量值。某些敏感器可以测量相同的量，在这种情况就得到简单的**冗余测量值**；另外一些敏感器可以测量仅仅与所需的过程间接有关的量。在这种多敏感器系统的范围内，估计的任务是处理敏感器的输出，使我们对感兴趣的过程获得“最好”的估计。而通过计算机实现的数据处理算法处理敏感器数据以提供要求的估值。这些估值可用来驱动显示器，也可作为被观测物理系统的控制信号，如图 1.0-3 所示。在现代多敏感器系统中，上述数据处理算法通常是用最优滤波理论推导出来的。

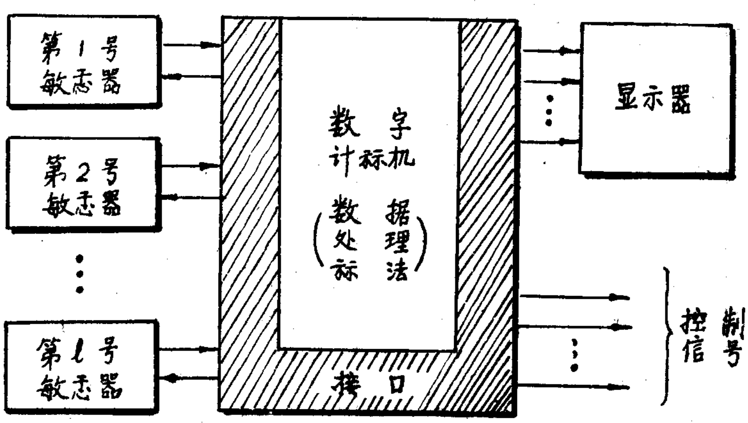

图 1.0-3 现代多敏感器系统

多敏感器系统的某些突出例子存在于导航领域中。外部测量值原来是用来以确定性方式修正导航变量的；例如，改变系统位置指示，使之与外部位置测量结果一致。这种作法忽略了两个重要的事实。首先，外部测量值本身包含着随机误差，这些误差与导航系统误差相比，可能是相当严重的误差。其次，导航系统误差主要是由随机的时变的导航敏感器误差引起的。最优地使用外部测量值和导航系统所提供的结果，可以获得比只有外部测量或导航系统时更好的导航精度。

在多敏感器导航系统中应用现代估计技术是从1960年年中开始的，那时最优递推滤波理论刚刚提出和发表。因为在一个典型的导航系统中误差的传播基本上是线性的，而且这些误差的线性组合可以从外部测量值探测到，所以应用卡尔曼滤波器对它们进行估计是很合适的。对于所有具有显著相关时间的系统误差，它也提供有用的估计。另外，卡尔曼滤波器还提供改善设计和工作的灵活性。作为一个时变滤波器，它也可以适应非平稳误差源，只要这些误差源的统计特性是已知的。在导航系统中结构的改变通过改变程序比较容易实现。卡尔曼滤波器为最优利用任意数目的外部测量值的组合和顺序

作了理论准备。它是一种有系统地应用所有可利用的外部测量值（不管它们的误差如何）以改善导航系统精度的技术。理论的这种应用在下面各章中将经常遇到。

在任何最优估计的实际应用中，主要的非硬件问题多半是**建立模型和实现性能设想**。显然，这些问题是有关的，而恰当地处理它们是使系统在实际环境中运行成功的必要条件。基于这种见解，显然可知：在任何工作系统中，一个估计理论的合理应用包括：**第一**，设计“最优”系统并对其特性进行计算机估价；**第二**，考虑价格限制、灵敏度特性、计算要求、测量程序等，设计一个合适的“亚优”系统；**第三**，构成并试验样机系统，并按要求作最后的调整或改变。

例 1.0-1

考虑一个包含两个敏感器的系统，每一个敏感器对一恒定**但未知**的量 x 进行一次测量 $z_i(i=1,\ 2)$，测量时存在随机独立、无偏的测量误差 $v_i(i=1,\ 2)$。设计一个数据处理算法，组合两个测量值以得出一个 x 的最优估值。

其测量值描述如下：

$$z_1=x+v_1 \quad 和 \quad z_2=x+v_2 \tag{1.0-1}$$

在没有别的任何信息的情况下，我们可以找到一个 x 的估值，它在形式上是测量值的线性函数（上标“∧”表示估值）

$$\hat{x}=k_1z_1+k_2z_2 \tag{1.0-2}$$

这里 k_1 和 k_2 是待定的系数。估值误差 $\tilde{x}$ 定义为

$$\tilde{x}=\hat{x}-x$$

我们希望取使 $\tilde{x}$ 的均方值最小作为最优判据，而且，我们要求我们选择的 k_1 和 k_2 与 x 值无关；如果估值是**无偏**的，这个条件*将满足，也就是说，如果

$$E[\tilde{x}]=E[k_1(x+v_1)+k_2(x+v_2)-x]=0 \tag{1.0-3}$$

k_1，k_2 将与 x 无关。这里 E 表示总体期望值**或平均值。根据期望的运算法则，在 $E[v_1]=E[v_2]=0$ 和 $E[x]=x$（因 x 是“非随机”的）的情况下，我们得到

$$k_2=1-k_1 \tag{1.0-4}$$

联合方程 (1.0-1) 到 (1.0-4)，计算得到均方误差为

$$E[\tilde{x}^2]=k_1^2\sigma_1^2+(1-k_1)^2\sigma_2^2 \tag{1.0-5}$$

这里 σ_1^2 和 σ_2^2 分别表示 v_1 和 v_2 的方差。将此式对 k_1 微分，并令结果等于零可得

$$2k_1\sigma_1^2-2(1-k_1)\sigma_2^2=0$$

或

$$k_1=\frac{\sigma_2^2}{\sigma_1^2+\sigma_2^2}$$

对应的最小均方估值误差是

$$E[\tilde{x}^2]=\left(\frac{1}{\sigma_1^2}+\frac{1}{\sigma_2^2}\right)^{-1} \tag{1.0-6}$$

---

* 加上这个条件，是因为 x 是未知的；因此，增益 k_1 和 k_2 必须是与 x 无关的。

** 数学期望运算在节 2.2中讨论。

由此可见，均方估值误差比任何一个测量的均方测量误差都要小。算法 (1.0-2)

$$\hat{x}=\left(\frac{\sigma_2^2}{\sigma_1^2+\sigma_2^2}\right)z_1+\left(\frac{\sigma_1^2}{\sigma_1^2+\sigma_2^2}\right)z_2 \qquad (1.0\text{-}7)$$

在各种感兴趣的极端情况下，有一定的意义——即，如果 $\sigma_1^2=\sigma_2^2$，则测量值得到平均；如果一个测量值是精确的（σ_1 或 σ_2 等于零），则另一个测量值就被丢弃；等等。

在例 1.0-1 中，我们研究的物理量是一个标量常数，而且只有这个常数的两个线性测量值是可得到的。一般，我们感兴趣的是在可以得到许多测量值的情况下，对时变的向量进行估计。

物理系统和测量值可按**静态**或**动态**、**连续**或**离散**、**线性**或**非线性**进行分类。此外，我们对实时数据处理算法(**滤波**和**予测**)和验后 (Post-experiment) 算法(**平滑**)最感兴趣。为了叙述简洁和理论上有力，在后面处理这些物理系统和数据处理算法时将用**状态空间**表示法来叙述，使用向量和矩阵的数学以及随机过程理论的概念。这些问题将在第二章和第三章内讨论。

习　题

题 1-1

当测量误差彼此相关时，考虑例 1.0-1 的问题；即 $E[v_1v_2]=\rho\sigma_1\sigma_2$，这里 ρ 是一相关系数（$|\rho|\leqslant 1$）。证明对最优估计，

$$k=\frac{\sigma_2^2-\rho\sigma_1\sigma_2}{\sigma_1^2+\sigma_2^2-2\rho\sigma_1\sigma_2}$$

及

$$E[\tilde{x}^2]=\frac{\sigma_1^2\sigma_2^2(1-\rho^2)}{\sigma_1^2+\sigma_2^2-2\rho\sigma_1\sigma_2}$$

从物理上解释当 $\rho=\pm1$ 时，$E[\tilde{x}^2]=0$ 的意义。

题 1-2

在 k_1 和 k_2 没有限制的情况下，从方程 (1.0-1) 和 (1.0-2) 计算 $E[\tilde{x}^2]$。并用 $E[x]=x$ 和 $E[x^2]=x^2$ 证明使 $E[\tilde{x}^2]$ 最小的增益值是 x 的函数。

题 1-3

考虑与例 1.0-1 类似的情况，但是在这里，可以得到的独立测量值是三个而不是两个。论证为什么估计器应当用下面的形式：

$$\hat{x}=k_1z_1+k_2z_2+(1-k_1-k_2)z_3$$

推导 k_1、k_2 的最优值，并用它们证明，对最优估计，有

$$E[\tilde{x}^2]=\left(\frac{1}{\sigma_1^2}+\frac{1}{\sigma_2^2}+\frac{1}{\sigma_3^2}\right)^{-1}$$

题 1-4

在试验期间溶液中物质的浓度按指数减小，在 t_1 和 t_2 时间对浓度进行含噪声的测

量，得到

$$z_i = x_0 e^{-t_i} + v_i \qquad (i=1,2)$$

希望得到初始浓度 x_0 的估值。说明无偏估值为

$$\hat{x}_0 = (ke^{t_1})z_1 + [(1-k)e^{t_2}]z_2$$

此处 k 是一个待定的常数。证明，使均方估值误差最小的 k 值是

$$k = \frac{\sigma_2^2}{\sigma_1^2 e^{-2(t_2-t_1)} + \sigma_2^2}$$

而对应的均方估值误差为

$$E[(\hat{x}_0 - x_0)^2] = \left(\frac{1}{\sigma_1^2}e^{-2t_1} + \frac{1}{\sigma_2^2}e^{-2t_2}\right)^{-1}$$

这里 σ_1^2 和 σ_2^2 是测量误差的方差。

参 考 文 献

1. Gauss, K. F., *Theoria Motus*, 1809; also, *Theory of the Motion of the Heavenly Bodies About the Sun in Conic Sections*, Dover, NewYork, 1963.

2. Fisher, R.A. "On an Absolute Criterion for Fitting Frequency Curves," *Messenger of Math.*, Vol. 41, 1912, p. 155.

3. Wiener, N., "The Extrapolation, Interpolation and Smoothing of Stationary Time Series," OSRD 370,Report to the Services 19, Research Project DIC-6037, MIT, February 1942.

4. Wiener, N. *The Extrapolation, Interpolation and Smoothing of Stationary Time Series*, John Wiley &Sons, Inc, New York, 1949.

5. Kolmogorov, A.N., "Interpolation and Extrapolation von Stationaren Zufalligen Folgen," *Bull. Acad.Sci. USSR*, Ser. Math. 5, 1941, pp. 3-14.

6. Zadeh, L.A., and Ragazzini, J.R., "An Extension of Wiener's Theory of Prediction," *J, Appl. Phys.* Vol. 21, July 1950, pp. 645-655,

7. Laning, J.H., Jr. and Battin, R.H., *Random Processes in Automatic Control*, McGraw-Hill Book Co., Inc., New york, 1956, Chapter 8.

8. Shinbrot, M., "Optimization Of Time-varying Linear Systems with Nonstationary Inputs," *Trans. ASME*, Vol. 80, 1958, pp. 457-462.

9. Swerling, P., "First Order Error Propagation in a Stagewise Smoothing Procedure for Satellite Observations," *J. Astronautical Sciences*, Vol. 6, 1959, pp. 46-52.

10. Kalman, R.E., "A New Approach to Linear Filtering and Prediction Problems," *J. Basic Eng.*, *March* 1960, pp. 35-46.

11. Kalman, R.E., and Bucy, R.S., "New Results in Linear Filtering and Prediction Theory," *J. Basic Eng.*, *March* 1961, pp. 95-108

12. Blum, M., "A Stagewise Parameter Estimation Procedure for Correlated Data," *Numer. Math*, Vol. 3. 1961, pp. 202-208.

13. Battin, R.H., "A Statistical Optimizing Navigation Procedure for Space Flight," *ARS Journal*, Vol. 32, 1962, pp. 1681-1696.

14. Sorenson, H.W., "Least-Squares Estimation, From Gauss to Kalman," *IEEE Spectrum*, July 1970, pp. 63-68.

第二章　基本数学方法的回顾

这一章回顾有关建立和应用现代估计理论所用到的数学方法。这些方法包括向量和矩阵运算以及它们在最小平方估计中的应用。还包括概率，随机变量和随机过程的介绍。后者对我们的工作很重要，因此给予较大的重视。本章的目的是为建立状态向量和协方差矩阵方法以及为它们今后的应用提供必需的数学工具。

注意，本章的取材既不是对这些课题的充分而详尽的讨论，也不是数学上严密的介绍。更详细的材料在引用的文献中可以找到。

2.1　向量、矩阵和最小平方

这一节包含与现代估计理论中应用的向量和矩阵方法有关的定义和运算。想要更充分了解这些内容的读者請参阅文献 1-3。

向　量　运　算

一组元素，排成一列，用下面有一横道的小写字母表示，叫做一个**向量**（更准确地说，一个列向量）。向量中元素的数目叫做向量的維数。这样，一个 n 維向量 $\underline{x}$ 是

$$\underline{x}=\begin{pmatrix}x_1\\x_2\\\vdots\\x_n\end{pmatrix}\qquad(2.1-1)$$

向量加法——两个向量的加法定义为

$$\underline{x}+\underline{y}\triangleq\begin{pmatrix}x_1+y_1\\x_2+y_2\\\vdots\\x_n+y_n\end{pmatrix}\qquad(2.1-2)$$

向量减法的定义类似。在向量加法和减法中，$\underline{x}$ 和 $\underline{y}$ 必须有相同的維数。

乘标量——一个向量可以乘以标量 k 得到

$$k\underline{x}=\begin{pmatrix}kx_1\\kx_2\\\vdots\\kx_n\end{pmatrix}\qquad(2.1-3)$$

零向量——每个元素为数 0 的向量。用 $\underline{0}$ 表示。

向量转置——一向量的转置定义如下：若 $\underline{x}$ 是列向量

$$\underline{x}=\begin{bmatrix}x_1\\x_2\\\vdots\\x_n\end{bmatrix} \tag{2.1-4}$$

它的转置是行向量

$$\underline{x}^T=[x_1,x_2\cdots x_n] \tag{2.1-5}$$

内积——量 $\underline{x}^T\underline{y}$ 称为内积或点积，其结果为**标量**

$$x^Ty=x_1y_1+x_2y_2+\cdots+x_ny_n \tag{2.1-6}$$

若 $\underline{x}^T\underline{y}=0$，则称 $\underline{x}$ 和 $\underline{y}$ **正交**。此外，向量 $\underline{x}$ 的长度平方是 $\underline{x}^T\underline{x}$，

$$\underline{x}^T\underline{x}=x_1^2+x_2^2+\cdots+x_n^2 \tag{2.1-7}$$

向量 $\underline{x}$ 的长度表示为

$$|\underline{x}|=\sqrt{\underline{x}^T\underline{y}} \tag{2.1-8}$$

外积——量 $\underline{x}\underline{y}^T$ 称为外积，其结果为一矩阵，

$$\underline{x}\underline{y}^T=\begin{pmatrix}x_1y_1 & x_1y_2 & \cdots & x_1y_n\\x_2y_1 & x_2y_2 & \cdots & x_2y_n\\\vdots & \vdots & & \vdots\\x_ny_1 & x_ny_2 & \cdots & x_ny_n\end{pmatrix}$$

类似地，我们可以组成矩阵 $\underline{x}\underline{x}^T$.

$$\underline{x}\underline{x}^T=\begin{pmatrix}x_1^2 & x_1x_2 & \cdots & x_1x_n\\x_2x_1 & x_2^2 & \cdots & x_2x_n\\\vdots & \vdots & & \vdots\\x_nx_1 & x_nx_2 & \cdots & x_n^2\end{pmatrix} \tag{2.1-9}$$

式中 $\underline{x}\underline{x}^T$ 称为向量 $\underline{x}$ 的**扩散矩阵**

向量导数——利用前面定义的运算，可以定义一向量的导数。对连续变化的向量 $\underline{x}(t)$ 和 $\underline{x}(t+\Delta t)$，我们得到

$$\underline{x}(t+\Delta t)-\underline{x}(t)=\begin{pmatrix}x_1(t+\Delta t)\\x_2(t+\Delta t)\\\vdots\\x_n(t+\Delta t)\end{pmatrix}-\begin{pmatrix}x_1(t)\\x_2(t)\\\vdots\\x_n(t)\end{pmatrix}=\begin{pmatrix}x_1(t+\Delta t)-x_1(t)\\x_2(t+\Delta t)-x_2(t)\\\vdots\\x_n(t+\Delta t)-x_n(t)\end{pmatrix} \tag{2.1-10}$$

用标量 $1/\Delta t$ 乘两边，取 $\Delta t\to 0$ 时的极限，得

$$\dot{\underline{x}}(t)=\begin{pmatrix}\dot{x}_1(t)\\\dot{x}_2(t)\\\vdots\\\dot{x}_n(t)\end{pmatrix} \tag{2.1-11}$$