【美】Dan Simon 著

OPTIMAL STATE ESTIMATION

KALMAN, $H\infty$, AND NONLINEAR APPROACHES

最优状态估计

——卡尔曼，$H\infty$及非线性滤波

张勇刚 李宁 奔粤阳 译

国防工业出版社
National Defense Industry Press

WILEY

最优状态估计
——卡尔曼，H_∞ 及非线性滤波

[美] Dan Simon 著
张勇刚 李宁 奔粤阳 译

国防工業出版社
·北京·

著作权合同登记　图字:军 - 2012 - 107 号

图书在版编目(CIP)数据

最优状态估计:卡尔曼,H_∞及非线性滤波/
(美)西蒙(Simon,D.)著;张勇刚,李宁,
奔粤阳译.—北京:国防工业出版社,2013.5
书名原文:Optimal state Kstimation-Kalman,
H_∞ and nonlinear approaches
ISBN 978-7-118-08808-3

Ⅰ.①最...　Ⅱ.①西...②张...③李...④奔
Ⅲ.①滤波器　Ⅳ.①TN713

中国版本图书馆 CIP 数据核字(2013)第 150684 号

Translation from the English Language Edition:
Optimal State Estimation:Kalman,H Infinity,and Nonlinear Approaches by Dan Simon.
ISBN 978 - 0 - 471 - 70858 - 2

※

国防工业出版社出版发行
(北京市海淀区紫竹院南路 23 号　邮政编码 100048)
三河市腾飞印务有限公司印刷
新华书店经售

*

开本 787 × 1092　1/16　**印张** 24¼　**字数** 555 千字
2013 年 5 月第 1 版第 1 次印刷　**印数** 1—2500 册　**定价** 68.00 元

(本书如有印装错误,我社负责调换)

国防书店:(010)88540777　　发行邮购:(010)88540776
发行传真:(010)88540755　　发行业务:(010)88540717

译者序

最优状态估计在很多领域都获得了广泛的应用。本书较全面地介绍了最优状态估计的方法,适合作为最优状态估计相关领域工程技术人员参考著作,也可作为相关课程的高年级本科生或研究生教材。本书分为四个部分。第1部分为基础知识,共包括四章。第1章回顾线性系统相关知识,建议读者在阅读本书之前具备线性系统理论的相关基础知识。第2章回顾概率论和随机过程相关内容。第3章介绍常量估计的最小二乘法及随机过程估计的维纳滤波,其中维纳滤波部分内容仅作简单介绍。第4章介绍状态的统计特性(均值和协方差)随时间的传播过程,第4章的内容连接本书的前三章和第2部分内容。

本书的第2部分介绍卡尔曼滤波。第5章介绍离散卡尔曼滤波的几种等价形式。第6章介绍卡尔曼滤波的其他形式,包括序贯卡尔曼滤波、信息滤波、平方根滤波和U-D滤波。第7章介绍一些卡尔曼滤波的扩展形式,包括相关噪声和有色噪声条件下的卡尔曼滤波,能够减少计算量的稳态滤波、衰减记忆滤波和带约束的卡尔曼滤波,这些能够使卡尔曼滤波的应用更加广泛。第8章介绍时间连续卡尔曼滤波。第9章介绍最优平滑,利用时间超前的量测信息去更好地估计当前状态。第10章介绍卡尔曼滤波的一些深入讨论,包括卡尔曼滤波的性能验证、降阶滤波器、鲁棒卡尔曼滤波,具有量测延迟和同步误差的卡尔曼滤波。

本书的第3部分介绍H_∞滤波。第11章介绍卡尔曼滤波的另一种等价形式作为推导H_∞滤波的基础,介绍了H_∞滤波。第12章讨论H_∞滤波的其他问题,包括混合卡尔曼/H_∞滤波和带约束的H_∞滤波。H_∞滤波还有很多问题有待学者深入研究。

本书的第4部分介绍非线性系统滤波方法。第13章介绍扩展卡尔曼滤波。第14章介绍无迹卡尔曼滤波,它具有相比于扩展卡尔曼滤波更加优越的性能。第15章介绍粒子滤波,这是解决非线性系统估计问题的最一般方法。无迹滤波和粒子滤波是相对较新的研究方向,有待学者继续深入研究。

由于本书译者水平有限,原稿中的部分错误已在翻译过程中纠正,但难免存在其他错误,请读者批评指正。

2012年12月

哈尔滨

目　录

第1部分　基础知识

第3部分 H_∞ 滤波

第4部分 非线性滤波

第1部分 基础知识

第1章 线性系统理论

> 最后,我们评论一下线性系统如此重要的原因。答案很简单:因为我们可以解出它们!
>
> ——Richard Feynman [Fey63, p. 25-4]

这一章回顾了线性系统的一些基本知识。在电气工程专业研究生的第一学期线性系统这门课中通常涵盖这些基本知识。最优状态估计理论主要依赖于矩阵论,包括矩阵运算,所以本章1.1节回顾了矩阵论。最优状态估计在线性和非线性系统中都能够得以应用,尽管状态估计在线性系统情况下的应用更为直接一些。1.2节简要回顾了线性系统,在1.3节讨论了非线性系统。状态空间系统可以在连续时间域或离散时间域中进行描述。通常,物理系统是在连续时间域中得以描述,而控制和状态估计算法则是在数字计算机上执行。1.4节讨论了可以得到连续时间系统的离散时间描述的方法。1.5节讨论了如何在数字计算机上仿真连续时间系统。1.6节和1.7节讨论了线性系统的稳定性、能控性以及能观性等基本概念。这些概念对于理解本书后面最优状态估计的内容是必要的。具有很强线性系统理论背景的同学可以跳过本章的内容。不过,在继续本书后面章节之前,通过回顾这些章节,至少可以有助于巩固状态估计的基本概念。

1.1 矩阵代数和矩阵运算

在这一节我们回顾一下矩阵、矩阵代数以及矩阵运算。这对于理解本书的其他部分是很必要的,因为最优状态估的计算法通常是以矩阵形式来描述的。

标量是一个单纯的数。例如,数字2是一个标量,数 $1+3j$ 是一个标量(我们在书中用 j 表示 -1 的平方根),数字 π 是一个标量。

向量由排列成行或列的标量所组成。例如,向量

$$\begin{bmatrix} 1 & 3 & \pi \end{bmatrix} \tag{1.1}$$

是一个3元向量,这个向量称为 1×3 向量,因为它有1行3列。这个向量也被称为一个行向量,因为它是按单独一行排列的。向量

$$\begin{bmatrix} -2 \\ \pi^2 \\ j \\ 0 \end{bmatrix} \tag{1.2}$$

是一个 4 元向量,这个向量也被称为 4×1 向量,因为它有 4 行 1 列。这个向量也称为一个列向量,因为它是按单独一列排列的。注意,一个数可以被视为一个一元向量,是一个简单的向量。

一个矩阵是由排列成矩形的标量所组成。例如,矩阵

$$\begin{bmatrix} -2 & 3 \\ 0 & \pi^2 \\ j & 0 \end{bmatrix} \tag{1.3}$$

是一个 3×2 矩阵,因为它有 3 行 2 列。在矩阵中,行数和列数都能够称为矩阵的维数,例如,前面方程中矩阵的维数是 3×2,注意向量可以被视为一个简单的矩阵,例如,式(1.1)是一个 1×3 矩阵。而数也能被视为一个简化的矩阵,例如,6 是一个 1×1 矩阵。

我们定义线性无关的行数为矩阵的秩,它也等于线性无关的列数。矩阵 A 的秩经常用符号 $\rho(A)$ 表示,矩阵 A 的秩总是小于或等于行数,它也总小于或等于列数。例如,矩阵

$$A = \begin{bmatrix} 1 & 2 \\ 2 & 4 \end{bmatrix} \tag{1.4}$$

的秩为 1,因为它只有一行线性无关,两行互为倍数。它也只有一列线性无关,其两列也互为倍数。

而矩阵

$$A = \begin{bmatrix} 1 & 3 \\ 2 & 4 \end{bmatrix} \tag{1.5}$$

的秩为 2,因为它有两行线性无关。也就是说,不存在非零数 c_1 和 c_2 使

$$c_1[1 \quad 3] + c_2[2 \quad 4] = [0 \quad 0] \tag{1.6}$$

所以这两行线性无关。它同样有两列线性无关,也就是说,不存在非零数 c_1 和 c_2 使

$$c_1\begin{bmatrix} 1 \\ 2 \end{bmatrix} + c_2\begin{bmatrix} 3 \\ 4 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \tag{1.7}$$

一个矩阵的所有组成元素都为零,那么它的秩为零。一个 $n \times m$ 矩阵的秩等于 n 和 m 中小的那个数,我们称之为满秩矩阵。一个 $n \times m$ 矩阵 A 的零度等于 $m - \rho(A)$。

可以通过将所有的行换成列来对矩阵(或向量)进行转置,也可以将所有列换成行。转置矩阵用上标 T 表示,例如,A^{T},如果 A 是 $r \times n$ 矩阵

$$A = \begin{bmatrix} A_{11} & \cdots & A_{1n} \\ \vdots & & \vdots \\ A_{r1} & \cdots & A_{rn} \end{bmatrix} \tag{1.8}$$

则 A^{T} 是 $n \times r$ 矩阵

$$A^{\mathrm{T}} = \begin{bmatrix} A_{11} & \cdots & A_{r1} \\ \vdots & & \vdots \\ A_{1n} & \cdots & A_{rn} \end{bmatrix} \tag{1.9}$$

注意，我们用符号 A_{ij} 去表示矩阵 A 中第 i 行和第 j 列所表示的数，对称矩阵满足 $A = A^{\mathrm{T}}$。

埃尔米特转置矩阵（或向量）是转置的复共轭矩阵。它用上标符号 H 表示，例如，A^{H}，如果

$$A = \begin{bmatrix} 1 & 2\mathrm{j} & 3-\mathrm{j} \\ 4\mathrm{j} & 5+\mathrm{j} & 1-3\mathrm{j} \end{bmatrix} \tag{1.10}$$

则

$$A^{\mathrm{H}} = \begin{bmatrix} 1 & -4\mathrm{j} \\ -2\mathrm{j} & 5-\mathrm{j} \\ 3+\mathrm{j} & 1+3\mathrm{j} \end{bmatrix} \tag{1.11}$$

埃尔米特矩阵满足 $A = A^{\mathrm{H}}$。

1.1.1 矩阵代数

矩阵加法和减法可以简单定义为对应元素的加法和减法。例如，

$$\begin{bmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{bmatrix} + \begin{bmatrix} 0 & 4 & 1 \\ 1 & -1 & -2 \end{bmatrix} = \begin{bmatrix} 1 & 6 & 4 \\ 4 & 1 & -1 \end{bmatrix} \tag{1.12}$$

$(A+B)$ 以及 $(A-B)$ 仅仅在 A 与 B 的维数相同时才有意义。

假设 A 是 $n \times r$ 矩阵，B 是 $r \times p$ 矩阵，A 与 B 的乘积写为 $C = AB$，矩阵乘积 C 中每一个元素的计算如下：

$$c_{ij} = \sum_{k=1}^{r} A_{ik} B_{kj} \quad i = 1,2,\cdots,n \quad j = 1,2,\cdots,p \tag{1.13}$$

矩阵乘积 AB 仅仅在 A 的列数与 B 的行数相同时才有意义。注意，相乘顺序不能调换。一般来说，$AB \neq BA$。

假设我们有 $n \times 1$ 向量 x。我们计算 1×1 乘积 $x^{\mathrm{T}} x$ 和 $n \times n$ 乘积 xx^{T} 如下：

$$x^{\mathrm{T}} x = [x_1 \cdots x_n] \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = x_1^2 + \cdots + x_n^2$$

$$xx^{\mathrm{T}} = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} [x_1 \cdots x_n] = \begin{bmatrix} x_1^2 & \cdots & x_1 x_n \\ \vdots & \ddots & \vdots \\ x_n x_1 & \cdots & x_n^2 \end{bmatrix} \tag{1.14}$$

假设我们有 $p \times n$ 矩阵 H 和 $n \times n$ 矩阵 P，则 H^{T} 是 $n \times p$ 矩阵。我们能计算 $p \times p$ 矩阵乘积 HPH^{T} 如下：

$$HPH^{\mathrm{T}} = \begin{bmatrix} H_{11} & \cdots & H_{1n} \\ \vdots & \ddots & \vdots \\ H_{p1} & \cdots & H_{pn} \end{bmatrix} \begin{bmatrix} P_{11} & \cdots & P_{1n} \\ \vdots & \ddots & \vdots \\ P_{n1} & \cdots & P_{nn} \end{bmatrix} \begin{bmatrix} H_{11} & \cdots & H_{p1} \\ \vdots & \ddots & \vdots \\ H_{1n} & \cdots & H_{pn} \end{bmatrix}$$

$$
= \begin{bmatrix} \sum_{j,k} H_{1j} P_{jk} H_{1k} & \cdots & \sum_{j,k} H_{1j} P_{jk} H_{pk} \\ \vdots & \ddots & \vdots \\ \sum_{j,k} H_{pj} P_{jk} H_{1k} & \cdots & \sum_{j,k} H_{pj} P_{jk} H_{pk} \end{bmatrix} \tag{1.15}
$$

这个和的矩阵可以写为矩阵的和,如下所示

$$
\begin{aligned}
HPH^{\mathrm{T}} &= \begin{bmatrix} H_{11} P_{11} H_{11} & \cdots & H_{11} P_{11} H_{p1} \\ \vdots & \ddots & \vdots \\ H_{p1} P_{11} H_{11} & \cdots & H_{p1} P_{11} H_{p1} \end{bmatrix} + \cdots + = \begin{bmatrix} H_{1n} P_{nn} H_{1n} & \cdots & H_{1n} P_{nn} H_{pn} \\ \vdots & \ddots & \vdots \\ H_{pn} P_{nn} H_{1n} & \cdots & H_{pn} P_{nn} H_{pn} \end{bmatrix} \\
&= H_1 P_{11} H_1^{\mathrm{T}} + \cdots + H_n P_{nn} H_n^{\mathrm{T}} \\
&= \sum_{j,k} H_j P_{jk} H_k^{\mathrm{T}}
\end{aligned} \tag{1.16}
$$

其中我们用符号 H_k 表示 H 的第 k 列。

我们不定义矩阵的除法。我们不能用一个矩阵去除另一个矩阵(除非分母矩阵是一个数)。

单位矩阵 I 是除对角线上的元素 1 以外全为 0 的方阵。例如,3×3 单位矩阵等于

$$
I = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \tag{1.17}
$$

单位矩阵满足特性 $AI = A$,且 $IA = A$(只要单位矩阵的维数是与矩阵 A 相匹配的),1×1 单位矩阵等于 1。对于方阵,矩阵的行列式可以被递归定义。数的行列式(一个 1×1 矩阵)等于该数本身。现在考虑 $n \times n$ 矩阵 A,用符号 $A^{(i,j)}$ 表示删除矩阵 A 的第 i 行、第 j 列后所形成的矩阵。我们定义 A 矩阵的行列式为

$$
|A| = \sum_{j=1}^{n} (-1)^{i+j} A_{ij} \, |A^{(i,j)}| \tag{1.18}
$$

其中 $i \in [1,n]$,这称为矩阵 A 沿其第 i 行的拉普拉斯展开式。我们看到可以从 $(n-1) \times (n-1)$ 矩阵的行列式的角度定义 $n \times n$ 矩阵 A 的行列式。相同的,可以从 $(n-2) \times (n-2)$ 矩阵的行列式的角度定义 $(n-1) \times (n-1)$ 矩阵的行列式。一直继续下去,直到从 1×1 矩阵的行列式的角度定义 2×2 矩阵的行列式为止,其中 1×1 矩阵的行列式是数。矩阵 A 的行列式也可以定义如下

$$
|A| = \sum_{i=1}^{n} (-1)^{i+j} A_{ij} \, |A^{(i,j)}| \tag{1.19}
$$

其中 $j \in [1,n]$,这被称为矩阵 A 沿其第 j 列的拉普拉斯展开式。有趣的是,式(1.18)(对任意 i 值)和式(1.19)(对任意 j 值)能得到相同的结果。从行列式的定义,我们可以知道

$$
\det[A_{11}] = A_{11}
$$

$$
\det\begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} = A_{11}A_{22} - A_{12}A_{21}
$$

$$\det\begin{bmatrix} A_{11} & A_{12} & A_{13} \\ A_{21} & A_{22} & A_{23} \\ A_{31} & A_{32} & A_{33} \end{bmatrix} = A_{11}(A_{22}A_{33} - A_{23}A_{32}) - A_{12}(A_{21}A_{33} - A_{23}A_{31}) + A_{13}(A_{21}A_{32} - A_{22}A_{31}) \tag{1.20}$$

假设 A 和 B 是方阵且具有相同的维数,行列式的一些有趣的特性如下

$$|AB| = |A||B| \tag{1.21}$$

且

$$|A| = \prod_{i=1}^{n} \lambda_i \tag{1.22}$$

其中 λ_i(A 的特征值)将在后面定义。

矩阵 A 的逆被定义为 A^{-1},满足 $AA^{-1} = A^{-1}A = I$。只有当矩阵 A 是方阵时才可逆。有些方阵没有逆,我们称之为奇异矩阵或者不可逆矩阵。在标量的情况下,没有逆的标量只有 0。但是在矩阵的情况下,很多矩阵都是奇异的。具有逆的矩阵,称之为非奇异矩阵或者可逆矩阵。例如

$$\begin{bmatrix} 1 & 0 \\ 2 & 3 \end{bmatrix}\begin{bmatrix} 1 & 0 \\ -2/3 & 1/3 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \tag{1.23}$$

因此,方程左边的两矩阵互为可逆矩阵。非奇异 $n \times n$ 矩阵 A 可以用很多等价的方式表述,其中如:

- A 是非奇异矩阵
- A^{-1} 存在
- A 的秩等于 n
- A 的行线性无关
- A 的列线性无关
- $|A| \neq 0$
- $Ax = b$ 对所有 b 有唯一 x 成立
- 0 不是 A 的特征值

定义对角线上元素的和为方阵的迹:

$$\mathrm{Tr}(A) = \sum_i A_{ii} \tag{1.24}$$

仅当矩阵为方阵时迹有意义。1×1 矩阵的迹等于其本身,也等于其标量值。方阵迹的一个有趣的特性是

$$\mathrm{Tr}(A) = \sum_i \lambda_i \tag{1.25}$$

也就是说,方阵的迹等于其特征值的和。一些有趣且有用的矩阵乘积特性如下:

$$\begin{aligned} (AB)^{\mathrm{T}} &= B^{\mathrm{T}}A^{\mathrm{T}} \\ (AB)^{-1} &= B^{-1}A^{-1} \\ \mathrm{Tr}(AB) &= \mathrm{Tr}(BA) \end{aligned} \tag{1.26}$$

上述式子成立的前提是假设矩阵的逆存在，且矩阵的维数匹配，矩阵乘法有意义。矩阵乘积的转置等于反序转置的乘积。矩阵乘积的逆等于反序逆的乘积。矩阵乘积的迹与矩阵相乘的顺序无关。

实数值列向量的 2 范数，我们也称为欧几里得范数，定义如下

$$\| x \|_2 = \sqrt{x^{\mathrm{T}} x} = \sqrt{x_1^2 + \cdots + x_n^2} \tag{1.27}$$

从式(1.14)我们知道

$$xx^{\mathrm{T}} = \begin{bmatrix} x_1^2 & \cdots & x_1 x_n \\ \vdots & \ddots & \vdots \\ x_n x_1 & \cdots & x_n^2 \end{bmatrix} \tag{1.28}$$

该矩阵的迹如下

$$Tr(xx^{\mathrm{T}}) = x_1^2 + \cdots + x_n^2 = \| x \|_2^2 \tag{1.29}$$

$n \times n$ 矩阵 A 有 n 个特征值和 n 个特征向量。如果下面方程成立，则 λ 是 A 的一个特征值，$n \times 1$ 向量 x 是 A 的一个特征向量

$$Ax = \lambda x \tag{1.30}$$

矩阵的特征值和特征向量都视为矩阵的特征量。$n \times n$ 矩阵有 n 个特征值，尽管有一些可能重复，这就像 n 阶多项式方程有 n 个根，尽管有些可能重复。从上面特征值和特征向量的定义我们可知：

$$Ax = \lambda x$$

$$A^2 x = A\lambda x = \lambda(Ax) = \lambda(\lambda x) = \lambda^2 x \tag{1.31}$$

所以，如果 A 有特征量(λ, x)，则 A^2 有特征量(λ^2, x)，当且仅当 A 中所有特征值不为零时，A^{-1}存在。如果 A 是对称矩阵，则它的所有特征值是实数。$n \times n$ 对称矩阵 A 可以按特性分为：正定、半正定、负定、半负定、不定矩阵。定义如下。

- 正定：对于所有 $n \times 1$ 非零向量 x，$x^{\mathrm{T}}Ax > 0$，这等价于 A 的所有特征值是正实数。如果 A 是正定，则 A^{-1}也是正定。
- 半正定：对于所有 $n \times 1$ 非零向量 x，$x^{\mathrm{T}}Ax \geqslant 0$，这等价于 A 的所有特征值是非负实数。半正定矩阵有时也称为非负定矩阵。
- 负定：对于所有 $n \times 1$ 非零向量 x，$x^{\mathrm{T}}Ax < 0$，这等价于 A 的所有特征值是负实数。如果 A 是负定，则 A^{-1}也是负定。
- 半负定：对于所有 $n \times 1$ 非零向量 x，$x^{\mathrm{T}}Ax \leqslant 0$，这等价于 A 的所有特征值是非正实数。半负定矩阵有时也称为非正定矩阵。
- 不定：不属于上述四种情况，也就是说，其特征值有些为正，有些为负。

有些书总结的正定性和负定性包括非对称矩阵。

$n \times 1$ 向量 x 的加权 2 范数定义为

$$\| x \|_Q^2 = \sqrt{x^{\mathrm{T}} Q x} \tag{1.32}$$

其中，Q 是正定矩阵。上述范数也被称为 x 的 Q 加权 2 范数。$x^{\mathrm{T}}Qx$ 形式的数值称为二次型，类似于方程中的二次项。

如果 A 是 $n\times m$ 矩阵，则其奇异值 σ 定义为

$$\sigma^2(A) = \lambda(A^{\mathrm{T}}A) = \lambda(AA^{\mathrm{T}}) \tag{1.33}$$

A 有 $\min(n,m)$ 个奇异值。AA^{T} 将有 n 个特征值，$A^{\mathrm{T}}A$ 将有 m 个特征值。如果 $n>m$，则 AA^{T} 的特征值将包含 $A^{\mathrm{T}}A$ 特征值加上 $n-m$ 个 0。这些附加的 0 都不被视为 A 的奇异值，因为 A 只有 $\min(n,m)$ 个奇异值。在奇异值计算过程中，这些知识可以帮助我们减少工作量。例如，如果 A 是 13×3 矩阵，则 3×3 矩阵 $A^{\mathrm{T}}A$ 要比 13×13 矩阵 AA^{T} 更加易于计算，都可以得到相同的三个奇异值。

1.1.2 矩阵逆引理

本节将推导矩阵逆引理，它在本书中将多次运用，在控制、估计理论以及信号处理等其他领域也将被经常用到。假设我们有分块矩阵 $\begin{bmatrix} A & B \\ C & D \end{bmatrix}$，其中 A 和 D 是可逆矩阵，B 和 C 可以是方阵也可以不是方阵。我们对 E 和 F 矩阵作如下定义：

$$\begin{aligned} E &= D - CA^{-1}B \\ F &= A - BD^{-1}C \end{aligned} \tag{1.34}$$

假设 E 是可逆矩阵，则我们可以得到

$$\begin{aligned} &\begin{bmatrix} A & B \\ C & D \end{bmatrix}\begin{bmatrix} A^{-1} + A^{-1}BE^{-1}CA^{-1} & -A^{-1}BE^{-1} \\ -E^{-1}CA^{-1} & E^{-1} \end{bmatrix} \\ &= \begin{bmatrix} I + BE^{-1}CA^{-1} - BE^{-1}CA^{-1} & -BE^{-1} + BE^{-1} \\ CA^{-1} + CA^{-1}BE^{-1}CA^{-1} - DE^{-1}CA^{-1} & -CA^{-1}BE^{-1} + DE^{-1} \end{bmatrix} \\ &= \begin{bmatrix} I & 0 \\ CA^{-1} - (D - CA^{-1}B)E^{-1}CA^{-1} & (D - CA^{-1}B)E^{-1} \end{bmatrix} \\ &= \begin{bmatrix} I & 0 \\ 0 & I \end{bmatrix} \end{aligned} \tag{1.35}$$

现在，我们假设 F 是可逆的，则我们可得

$$\begin{aligned} &\begin{bmatrix} A & B \\ C & D \end{bmatrix}\begin{bmatrix} F^{-1} & -A^{-1}BE^{-1} \\ -D^{-1}CF^{-1} & E^{-1} \end{bmatrix} \\ &= \begin{bmatrix} AF^{-1} - BD^{-1}CF^{-1} & -BE^{-1} + BE^{-1} \\ CF^{-1} - CF^{-1} & -CA^{-1}BE^{-1} + DE^{-1} \end{bmatrix} \\ &= \begin{bmatrix} (A - BD^{-1}C)F^{-1} & 0 \\ 0 & (D - CA^{-1}B)E^{-1} \end{bmatrix} \\ &= \begin{bmatrix} I & 0 \\ 0 & I \end{bmatrix} \end{aligned} \tag{1.36}$$

式(1.35)和式(1.36)是 $\begin{bmatrix} A & B \\ C & D \end{bmatrix}$ 逆的两个表达式。因为这两个是同一个矩阵逆的表达式，它们应该相等。因此，我们得出结论，矩阵左上方的分块阵是相等的，即

$$F^{-1} = A^{-1} + A^{-1}BE^{-1}CA^{-1} \tag{1.37}$$