Pure Mathematics 理论数学, 2024, 14(4), 53-57
Published Online April 2024 in Hans. https://www.hanspub.org/journal/pm
https://doi.org/10.12677/pm.2024.144110

Hans汉斯

Sherman-Morrison公式及其应用

梁国宏，冯军庆，宋修朝

空军工程大学基础部，陕西 西安

收稿日期：2024年2月2日；录用日期：2024年3月1日；发布日期：2024年4月16日

摘 要

Sherman-Morrison公式是求矩阵之和的逆矩阵的一种特殊方法，在最优化BFGS算法和循环三对角线性方程组的求解等方面有着重要的应用。

关键词

Sherman-Morrison公式逆矩阵，BFGS算法，循环三对角线性方程组的求解

Sherman-Morrison Formula and Its Application

Guohong Liang, Junqing Feng, Xiuchao Song

Fundamentals Department, Air Force Engineering University, Xi'an Shaanxi

Received: Feb. 2nd, 2024; accepted: Mar. 1st, 2024; published: Apr. 16th, 2024

Abstract

The Sherman Morrison formula is a special method for finding the inverse matrix of the sum of matrices. It has important applications in optimizing the BFGS algorithm and solving cyclic tridiagonal linear systems of equations.

Keywords

Sherman-Morrison Formula Inverse Matrix, BFGS Algorithm, The Solution of Recurrent Tridiagonal Linear Equation System

文章引用: 梁国宏, 冯军庆, 宋修朝. Sherman-Morrison 公式及其应用[J]. 理论数学, 2024, 14(4): 53-57.
DOI: 10.12677/pm.2024.144110

1. 引言

Sherman-Morrison 公式是一种用于求解线性方程组描述的方法，以 Jack Sherman 和 Winifred J. Morrison 命名的，可用于计算机科学领域，特别是通信、网络和信号处理中。

由于矩阵之和的逆不一定可逆，即使可逆，也没有统一的公式，Sherman-Morrison 公式[1] [2]在线性代数中用来求解逆矩阵的一种特殊方法，可逆矩阵与某个列乘行的矩阵之和是一定可逆的，并用公式直接得到逆矩阵。在最优化 BFGS 算法和循环三对角线性方程组的求解等方面有着重要的应用。

可以用于增量计算矩阵的逆，从而节省存储空间和计算时间。比如，当需要计算一个大型矩阵的逆时，对矩阵逐步添加向量，从而逐步计算出矩阵的逆。

设 $A\in R^{n\times n}$ 为可逆方阵，$u,v\in R^{n}$ 为列向量，$1+v^{\mathrm{T}}A^{-1}u\neq 0$，需要求解线性方程 $\left(A+uv^{\mathrm{T}}\right)x=b$，考虑求解两个方程：$Ay=b, Az=u$，有 $y=A^{-1}b, z=A^{-1}u$，得

$$x+zv^{\mathrm{T}}x=y$$

注意到 $v^{\mathrm{T}}x$ 是一个数，令 $\alpha=v^{\mathrm{T}}x$，有 $x+z\alpha=y$，两端左乘 v^{T}，即 $v^{\mathrm{T}}x+v^{\mathrm{T}}z\alpha=v^{\mathrm{T}}y$，故 $\alpha+v^{\mathrm{T}}z\alpha=v^{\mathrm{T}}y$，得

$$\alpha=\frac{v^{\mathrm{T}}y}{1+v^{\mathrm{T}}z}=\frac{v^{\mathrm{T}}A^{-1}b}{1+v^{\mathrm{T}}A^{-1}u}$$

有

$$x=y-z\alpha=A^{-1}b-A^{-1}u\frac{v^{\mathrm{T}}A^{-1}b}{1+v^{\mathrm{T}}A^{-1}u}$$

实际上，Sherman-Morrison 公式在解线性方程有着重要的应用，即通过增加列向量来求解线性方程。设 $A\in R^{n\times n}$ 为可逆方阵，B 是一个 n 维列向量，考虑线性方程

$$AX=B$$

解为

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

令 $B'=\left(B,u\right)$，即在 B 的右侧增加一个 n 维列向量 u，则线性方程变为

$$AX'=B'$$

根据伴随矩阵，解为

$$X'=A^{-1}B'=A^{-1}\left(B,u\right)=\left(A+uv^{\mathrm{T}}\right)^{-1}\left(B+u\alpha\right)$$

其中 $\alpha=-\dfrac{v^{\mathrm{T}}A^{-1}u}{1+v^{\mathrm{T}}A^{-1}u}$。将解 X' 表示为

$$X'=A^{-1}B-\frac{A^{-1}uv^{\mathrm{T}}A^{-1}}{1+v^{\mathrm{T}}A^{-1}u}A^{-1}B$$

于是

$$X'=\left(A^{-1}-\frac{A^{-1}uv^{\mathrm{T}}A^{-1}}{1+v^{\mathrm{T}}A^{-1}u}\right)B$$

可见，通过增加列向量 u 来解线性方程的解法。

2. Sherman-Morrison 公式

设 $A\in R^{n\times n}$ 为可逆方阵，$u,v\in R^{n}$ 为列向量，则 $A+uv^{\mathrm{T}}$ 可逆充要条件为 $1+v^{\mathrm{T}}A^{-1}u\neq 0$，且当 $A+uv^{\mathrm{T}}$ 可逆时，有[3] [4]

$$\left(A+uv^{\mathrm{T}}\right)^{-1}=A^{-1}-\frac{A^{-1}uv^{\mathrm{T}}A^{-1}}{1+v^{\mathrm{T}}A^{-1}u}$$

证明(充分性)当 $1+v^{\mathrm{T}}A^{-1}u\neq 0$ 时，

$$\begin{aligned}\left(A+uv^{\mathrm{T}}\right)\left(A^{-1}-\frac{A^{-1}uv^{\mathrm{T}}A^{-1}}{1+v^{\mathrm{T}}A^{-1}u}\right)&=AA^{-1}+uv^{\mathrm{T}}A^{-1}-\frac{AA^{-1}uv^{\mathrm{T}}A^{-1}+uv^{\mathrm{T}}A^{-1}uv^{\mathrm{T}}A^{-1}}{1+v^{\mathrm{T}}A^{-1}u}\\&=E+uv^{\mathrm{T}}A^{-1}-\frac{uv^{\mathrm{T}}A^{-1}+uv^{\mathrm{T}}A^{-1}uv^{\mathrm{T}}A^{-1}}{1+v^{\mathrm{T}}A^{-1}u}\\&=E+uv^{\mathrm{T}}A^{-1}-\frac{u\left(1+v^{\mathrm{T}}A^{-1}u\right)v^{\mathrm{T}}A^{-1}}{1+v^{\mathrm{T}}A^{-1}u}\\&=E+uv^{\mathrm{T}}A^{-1}-uv^{\mathrm{T}}A^{-1}\\&=E\end{aligned}$$

因此，当 $1+v^{\mathrm{T}}A^{-1}u\neq 0$ 时，有

$$\left(A+uv^{\mathrm{T}}\right)^{-1}=A^{-1}-\frac{A^{-1}uv^{\mathrm{T}}A^{-1}}{1+v^{\mathrm{T}}A^{-1}u}$$

(必要性)当 $u=0$ 时，显然有 $1+v^{\mathrm{T}}A^{-1}u=1\neq 0$。当 $u\neq 0$ 时，用反证法证明该命题成立。假设 $A+uv^{\mathrm{T}}$ 可逆时，但 $1+v^{\mathrm{T}}A^{-1}u=0$ 时，有

$$\left(A+uv^{\mathrm{T}}\right)A^{-1}u=u+u\left(v^{\mathrm{T}}A^{-1}u\right)=u\left(1+v^{\mathrm{T}}A^{-1}u\right)=0$$

因为 $A+uv^{\mathrm{T}}$ 可逆，故 $A^{-1}u=0$，又因为 A^{-1} 可逆，故 $u=0$，此与假设 $u\neq 0$ 矛盾。因此，当 $A+uv^{\mathrm{T}}$ 可逆时，但 $1+v^{\mathrm{T}}A^{-1}u\neq 0$ 时，有 $1+v^{\mathrm{T}}A^{-1}u\neq 0$。

3. Sherman-Morrison 公式的应用

1) 当 $A=E_n$ 时的 Sherman-Morrison 公式

在 Sherman-Morrison 公式中，令 $A=E_n$，则有 $E+uv^{\mathrm{T}}$ 可逆充要条件为 $1+v^{\mathrm{T}}u\neq 0$，且当 $E+uv^{\mathrm{T}}$ 可逆时，有

$$\left(E+uv^{\mathrm{T}}\right)^{-1}=E-\frac{uv^{\mathrm{T}}}{1+v^{\mathrm{T}}u}$$

再令 $v=u$，则 $1+u^{\mathrm{T}}u>0$，因此，$E+u^{\mathrm{T}}u$ 可逆，且

$$\left(E+uu^{\mathrm{T}}\right)^{-1}=E-\frac{uu^{\mathrm{T}}}{1+u^{\mathrm{T}}u}$$

2) Sherman-Morrison 公式在 BFGS 算法中的应用

Sherman-Morrison 公式在 BFGS 算法中的应用，可用来求解 BFGS 算法中近似 Hessian 矩阵的逆。

在 BFGS 算法中，得到递推公式

$$B_{k+1}=B_k+\frac{q_k q_k^{\mathrm{T}}}{q_k^{\mathrm{T}} p_k}-\frac{B_k p_k p_k^{\mathrm{T}} B_k}{p_k^{\mathrm{T}} B_k p_k}$$

其中，$p_k=x^{(k+1)}-x^{(k)}$，$q_k=\nabla f\left(x^{(k+1)}\right)-\nabla f\left(x^{(k)}\right)$。

令 $H=B_k+\frac{q_k q_k^{\mathrm{T}}}{q_k^{\mathrm{T}} p_k}$，注意到 $\frac{B_k p_k p_k^{\mathrm{T}} B_k}{p_k^{\mathrm{T}} B_k p_k}=\frac{1}{p_k^{\mathrm{T}} B_k p_k}\left(B_k p_k\right)\left(B_k p_k\right)^{\mathrm{T}}$，其中 $p_k^{\mathrm{T}} B_k p_k$ 为一个数，且 $B_k p_k \in R^n$，因此利用 Sherman-Morrison 公式，有

$$\begin{aligned}
B_{k+1}^{-1}&=\left(H-\frac{1}{p_k^{\mathrm{T}} B_k p_k}\left(B_k p_k\right)\left(B_k p_k\right)^{\mathrm{T}}\right)^{-1}\\
&=H^{-1}-\frac{H^{-1}\left(-\frac{1}{p_k^{\mathrm{T}} B_k p_k}\left(B_k p_k\right)\left(B_k p_k\right)^{\mathrm{T}}\right) H^{-1}}{1+\left(-\frac{1}{p_k^{\mathrm{T}} B_k p_k}\right)\left(B_k p_k\right)^{\mathrm{T}} H^{-1}\left(B_k p_k\right)}\\
&=H^{-1}+H^{-1} \frac{B_k p_k p_k^{\mathrm{T}} B_k}{p_k^{\mathrm{T}} B_k p_k-p_k^{\mathrm{T}} B_k H^{-1} B_k p_k} H^{-1}
\end{aligned}$$

对于 H^{-1}，再次利用 Sherman-Morrison 公式，有

$$\begin{aligned}
H^{-1}&=\left(B_k+\frac{q_k q_k^{\mathrm{T}}}{q_k^{\mathrm{T}} p_k}\right)^{-1}=B_k^{-1}-\frac{B_k^{-1} \frac{q_k q_k^{\mathrm{T}}}{q_k^{\mathrm{T}} p_k} B_k^{-1}}{1+\frac{1}{q_k^{\mathrm{T}} p_k} q_k^{\mathrm{T}} B_k^{-1} q_k}\\
&=B_k^{-1}-\frac{B_k^{-1} \frac{q_k q_k^{\mathrm{T}}}{q_k^{\mathrm{T}} p_k} B_k^{-1}}{1+\frac{1}{q_k^{\mathrm{T}} p_k} q_k^{\mathrm{T}} B_k^{-1} q_k}=B_k^{-1}-\frac{B_k^{-1} q_k q_k^{\mathrm{T}} B_k^{-1}}{q_k^{\mathrm{T}} p_k+q_k^{\mathrm{T}} B_k^{-1} q_k}
\end{aligned}$$

将 H^{-1} 的表达式代回 B_{k+1}^{-1} 中，得到

$$B_{k+1}^{-1}=\left(I-\frac{p_k q_k^{\mathrm{T}}}{p_k^{\mathrm{T}} q_k}\right) B_k^{-1}\left(I-\frac{p_k q_k^{\mathrm{T}}}{p_k^{\mathrm{T}} q_k}\right)+\frac{p_k p_k^{\mathrm{T}}}{p_k^{\mathrm{T}} q_k}$$

3) Sherman-Morrison 公式在循环三对角线性方程组的求解中的应用

下面介绍循环三对角线性方程组的求解。所谓循环三对角线性方程组，指的是系数矩阵为如下形式：

$$A=\begin{bmatrix}
b_1 & c_1 & 0 & \cdots & 0 & a_1\\
a_2 & b_2 & c_2 & 0 & \vdots & 0\\
0 & \ddots & \ddots & \ddots & 0 & \vdots\\
\vdots & \vdots & a_{n-2} & b_{n-2} & c_{n-2} & 0\\
0 & \cdots & \cdots & a_{n-1} & b_{n-1} & c_{n-1}\\
c_n & 0 & \cdots & 0 & a_n & b_n
\end{bmatrix}$$

循环三对角线性方程组可写成 $Ax=d$，其中 $d=\left(d_1,d_2,\cdots,d_n\right)^{\mathrm{T}}$。

令 $u=\left(\lambda,0,0,\cdots,c_n\right)^{\mathrm{T}}$，$v=\left(1,0,0,\cdots,\frac{a_1}{\lambda}\right)^{\mathrm{T}}$，且 $A=A'+uv^{\mathrm{T}}$，其中

$$A' = \begin{bmatrix} b_1 - \lambda & c_1 & 0 & \cdots & 0 & 0 \\ a_2 & b_2 & c_2 & 0 & \vdots & 0 \\ 0 & \ddots & \ddots & \ddots & 0 & \vdots \\ \vdots & \vdots & a_{n-2} & b_{n-2} & c_{n-2} & 0 \\ 0 & \cdots & \cdots & a_{n-1} & b_{n-1} & c_{n-1} \\ c_n & 0 & \cdots & 0 & a_n & b_n - \dfrac{a_1 c_n}{\lambda} \end{bmatrix}$$

A' 为三对角矩阵。根据以上的理论知识，只需求解以下两个方程

$$A'y = d, A'z = u$$

这样就能根据 y, z 求出 x。

总之，Sherman-Morrison 公式给出了求矩阵之和的逆矩阵的一种特殊方法，在矩阵论中有着重要的应用，除此之外，在最优化 BFGS 算法和循环三对角线性方程组的求解等方面有着重要的应用。

参考文献

[1] 陈宝林. 最优化理论与算法[M]. 北京: 清华大学出版社, 2005.

[2] 王景恒. 最优化理论与方法[M]. 北京: 北京理工大学出版社, 2019.

[3] 袁亚湘, 孙文瑜. 最优化理论与方法[M]. 北京: 科学出版社, 1997.

[4] Auslender, A. and Teboulle, M. (2003) Asymptotic Cones and Functions in Optimization and Variational Inequalities. Springer, New York.