三十分钟理解：矩阵Cholesky分解，及其在求解线性方程组、矩阵逆的应用

写一篇关于Cholesky分解的文章，作为学习笔记，尽量一文看懂矩阵Cholesky分解，以及用Cholesky分解来求解对称正定 线性 方程组，以及求对称
的逆的应用。

文章目录

先简单理解下正定矩阵和半正定矩阵的定义[1][2][3]：

- 给定一个 $\mathrm{n} \times \mathrm{n}$ 的实对称矩阵 A ，若对于任意长度为 n 的非零向量 X，有 $\mathrm{X}^{\mathrm{T}}\mathrm{AX} > 0$ 恒成立，则矩阵 A 是一个正定矩阵。
- 给定一个 $\mathrm{n} \times \mathrm{n}$ 的实对称矩阵 A ，若对于任意长度为 n 的非零向量 X，有 $\mathrm{X}^{\mathrm{T}}\mathrm{AX} \geq 0$ 恒成立，则矩阵 A 是一个半正定矩阵。

等价命题

对于n阶实对称矩阵A，下列条件是等价的：

（1）A是正定矩阵；

（2）A的一切顺序主子式均为正；

（3）A的一切主子式均为正；

（4）A的特征值均为正；

（5）存在实可逆矩阵C，使A=C'C；

（6）存在秩为n的m×n实矩阵B，使A=B'B；

（7）存在主对角线元素全为正的实三角矩阵R，使A=R'R

直接Cholesky分解

定理——若$\mathrm{A} \in \mathrm{R}^{\mathrm{n}\times\mathrm{n}}$是对称 正定矩阵 ，则存在一个对角元全为正数的下三角矩阵$\mathrm{L} \in \mathrm{R}^{\mathrm{n}\times\mathrm{n}}$，使得$\mathrm{A} = \mathrm{LL}^{\mathrm{T}}$成立。

L是一个下三角形，形式是这样的：

$$L = \begin{pmatrix} l_{11} & 0 & \cdots & 0 \\ l_{21} & l_{22} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ l_{n1} & l_{n2} & \cdots & l_{nn} \end{pmatrix} \qquad L^T = \begin{pmatrix} l_{11} & l_{21} & \cdots & l_{n1} \\ 0 & l_{22} & \cdots & l_{n2} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & l_{nn} \end{pmatrix}$$

推导$\mathrm{A} = \mathrm{LL}^{\mathrm{T}}$： 我们先令

$$\mathrm{A} = \begin{bmatrix} \mathrm{a}_{11} & \mathrm{A}_{21}^{\mathrm{T}} \\ \mathrm{A}_{21} & \mathrm{A}_{22} \end{bmatrix}, \quad \mathrm{L} = \begin{bmatrix} \mathrm{l}_{11} & 0 \\ \mathrm{L}_{21} & \mathrm{L}_{22} \end{bmatrix}, \quad \mathrm{L}^{\mathrm{T}} = \begin{bmatrix} \mathrm{l}_{11} & \mathrm{L}_{21}^{\mathrm{T}} \\ 0 & \mathrm{L}_{22}^{\mathrm{T}} \end{bmatrix}$$

其中$\mathrm{a_{11}}$和$\mathrm{l_{11}}$是一个标量，$\mathrm{A_{21}}$和$\mathrm{L_{21}}$是一个列向量，$\mathrm{A_{22}}$是一个n-1阶的方阵，而$\mathrm{L_{22}}$是一个n-1阶的下三角形。那么：

$$\begin{bmatrix} \mathrm{a_{11}} & \mathrm{A_{21}^T} \\ \mathrm{A_{21}} & \mathrm{A_{22}} \end{bmatrix} = \begin{bmatrix} \mathrm{l_{11}} & 0 \\ \mathrm{L_{21}} & \mathrm{L_{22}} \end{bmatrix} \begin{bmatrix} \mathrm{l_{11}} & \mathrm{L_{21}^T} \\ 0 & \mathrm{L_{22}^T} \end{bmatrix} = \begin{bmatrix} \mathrm{l_{11}^2} & \mathrm{l_{11}L_{21}^T} \\ \mathrm{l_{11}L_{21}} & \mathrm{L_{21}L_{21}^T + L_{22}L_{22}^T} \end{bmatrix}$$

未知量只有标量$\mathrm{l_{11}}$，列向量$\mathrm{L_{21}}$，和下三角形$\mathrm{L_{22}}$，也是我们要求的。很容易得到：

$$\mathrm{l_{11}} = \sqrt{\mathrm{a_{11}}}$$

$$\mathrm{L_{21}} = \frac{1}{\mathrm{l_{11}}}\mathrm{A_{21}}$$

$$\mathrm{L_{22}L_{22}^T} = \mathrm{A_{22}} - \mathrm{L_{21}L_{21}^T}$$

其中$\mathrm{l_{11}}$，$\mathrm{L_{21}}$我们直接可以求出来了，并且可以求出$\mathrm{A'_{22}} = \mathrm{A_{22}} - \mathrm{L_{21}L_{21}^T}$。

而$\mathrm{A'_{22}} = \mathrm{L_{22}L_{22}^T}$又是一个Cholesky分解！被分解的矩阵是一个n-1阶方阵$\mathrm{A'_{22}}$。因此，Cholesky分解算法具有递归性质，每一轮可以求出L的一列，
求，就可以把整个L求出来。

一个例子[4]：

设：

$$A = \begin{bmatrix} 25 & 15 & -5 \\ 15 & 18 & 0 \\ -5 & 0 & 11 \end{bmatrix} = \begin{bmatrix} l_{11} & 0 & 0 \\ l_{21} & l_{22} & 0 \\ l_{31} & l_{32} & l_{33} \end{bmatrix} \begin{bmatrix} l_{11} & l_{21} & l_{31} \\ 0 & l_{22} & l_{32} \\ 0 & 0 & l_{33} \end{bmatrix}$$

根据上文的公式，有：

$$l_{11} = \sqrt{a_{11}} = 5$$

$$L_{21} = \frac{1}{l_{11}}A_{21} = \frac{1}{5}\begin{bmatrix} 15 \\ -5 \end{bmatrix} = \begin{bmatrix} 3 \\ -1 \end{bmatrix}$$

$$A_{22} - L_{21}L_{21}^T = L_{22}L_{22}^T$$

$$A_{22} - L_{21}L_{21}^T = L_{22}L_{22}^T$$

$$\begin{bmatrix} 18 & 0 \\ 0 & 11 \end{bmatrix} - \begin{bmatrix} 3 \\ -1 \end{bmatrix}\begin{bmatrix} 3 & -1 \end{bmatrix} = \begin{bmatrix} l_{22} & 0 \\ l_{32} & l_{33} \end{bmatrix}\begin{bmatrix} l_{22} & l_{32} \\ 0 & l_{33} \end{bmatrix}$$

$$\begin{bmatrix} 9 & 3 \\ 3 & 10 \end{bmatrix} = \begin{bmatrix} l_{22} & 0 \\ l_{32} & l_{33} \end{bmatrix}\begin{bmatrix} l_{22} & l_{32} \\ 0 & l_{33} \end{bmatrix}$$

(注意，这里已经是n-1阶的Cholesky分解)

$$l_{22} = \sqrt{9} = 3$$

$$l_{32} = \frac{1}{3}3 = 1$$

$$10 = l_{32}^2 + l_{33}^2 = 1 + l_{33}^2$$

$$l_{33} = \sqrt{10 - 1} = 3$$

综上：

$$A = \begin{bmatrix} 25 & 15 & -5 \\ 15 & 18 & 0 \\ -5 & 0 & 11 \end{bmatrix} = \begin{bmatrix} 5 & 0 & 0 \\ 3 & 3 & 0 \\ -1 & 1 & 3 \end{bmatrix}\begin{bmatrix} 5 & 3 & -1 \\ 0 & 3 & 1 \\ 0 & 0 & 3 \end{bmatrix}$$

另外，上述的方法需要进行开方，这有可能损失精度和增加运算量，为了避免开方，Cholesky分解有个改进的版本。将对称正定矩阵通过分解成$\mathrm{A} =$
其中L是单位下三角矩阵（单位下三角矩阵的对角线右上方的系数全部为零,左下方的系数全为一），D是对角均为正数的对角矩阵。把这一分解叫做LD
Cholesky分解的变形。具体先不展开了，可以参考[5]，以及其中的参考代码。

分块Cholesky分解

这一部分其实原理和上面是一样的，只是$\mathrm{l_{11}}$也是一个矩阵块来做，这样算法就是在一个矩阵块粒度递归往下做，效率上快很多，可以比较容易利用到
加速。我主要参考了文献[6]的内容，只做一个笔记用。
首先，把A矩阵分块，其中$\mathrm{A_{11}}$是一个$\mathrm{r \times r}$方阵，r是我们设定的可以采用直接cholesky分解算法求解的块大小。

$$A=\left[\begin{array}{c|c} A_{11} & B^{\mathrm{T}} \\ \hline B & \hat{A} \end{array}\right] \quad A_{11}\in\Re^{r\times r}$$

$$A=\left[\begin{array}{c|c} A_{11} & B^{\mathrm{T}} \\ \hline B & \hat{A} \end{array}\right]=\left[\begin{array}{c|c} L_{11} & 0 \\ \hline S & \hat{L} \end{array}\right]\left[\begin{array}{c|c} L_{11}^{\mathrm{T}} & S^{\mathrm{T}} \\ \hline 0 & \hat{L}^{\mathrm{T}} \end{array}\right]$$

类似的，位面可以得到：

$$L_{11}=cholesky(A_{11}) \tag{7}$$

$$S=B\cdot L_{11}^{-\mathrm{T}} \tag{8}$$

$$\hat{L}\cdot\hat{L}^{\mathrm{T}}=\hat{A}-S\cdot S^{\mathrm{T}} \tag{9}$$

其中公式（8）其实我们一般计算的是线性 方程组 ：

$$\mathrm{S}\cdot\mathrm{L}_{11}^{\mathrm{T}}=\mathrm{B}$$

其中$\mathrm{L}_{11}^{\mathrm{T}}$是一个上三角形，因此，我们可以比较容易求出S（先直接可以求出S的第一列，然后是第二列，以此类推）。可以直接调用各种BLAS库的trs
[7]。

公式（9）又是一个Cholesky分解，一直递归下去采用分块Cholesky分解，直到$\hat{\mathrm{A}}-\mathrm{S}\cdot\mathrm{S}^{\mathrm{T}}$的size小于等于$\mathrm{r}\times\mathrm{r}$就不在分解了，最后采用一次直接C
解。

示意图以及流程总结如下：

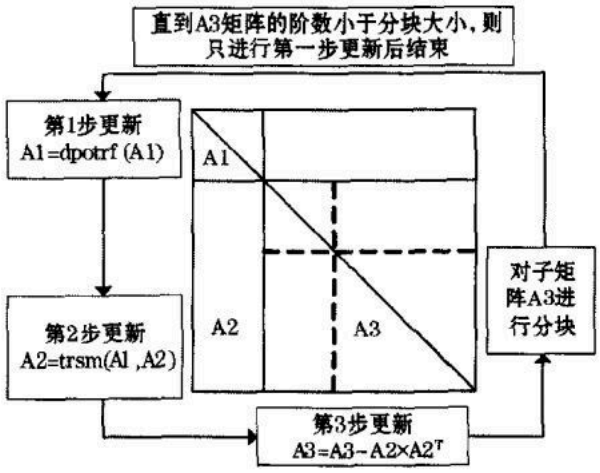

图 1　基于块的 Cholesky 分解

基于块的 Cholesky 分解算法如下：

Step1 使用 Intel MKL 库中的 dpotrf 函数计算第一个子块 A1 的 Cholesky 分解；

Step2 使用 BLAS 库中的 Level 3 函数 *trsm* 求解矩阵方程 $X \times A1^T = A2$；

Step3 使用 BLAS 库中的 Level 3 函数 *syrk* 更新尾部矩阵 $A3 = A3 - A2 \times A2^T$；

Step4 对尾部矩阵 A3 重新分块并针对 A3 重复执行以上三步，直到计算完最后一个子块的 Cholesky 分解则结束。

在余部矩阵 A3 的阶数较大时，Step 3 占据一次循环约 80% 的计算时间。而 A1 由于分块大小的固定，计算量不会改变，Step 1 所占的计算时间非常小。当 A2 的行远大于分块值时，Step 2 几乎占据了剩余 20% 的计算时间。

到这里就把Cholesky分解的计算方法讲清楚了。

Cholesky分解应用于求解线性方程组，以及矩阵求逆

如果A为对称正定矩阵，现在要求解线性方程组$AX = B$：
步骤：

1. 求A的Cholesky分解，得到$A = LL^T$
2. 把AX看成$L(L^T X) = LY$，求$LY = B$，得到Y
3. 求$L^T X = Y$，得到X

这样就求出了线性方程组$AX = B$的解X。

类似的，如果A为对称正定矩阵，我们要求A^{-1}，我们实际上只要求解线性方程组$AX = I$：
步骤：

1. 求A的Cholesky分解，得到$A = LL^T$
2. 把AX看成$L(L^T X) = LY$，求$LY = I$，得到Y
3. 求$L^T X = Y$，得到X

这样就求出了逆矩阵$X = A^{-1}$。

参考资料

[1] https://www.jiqizhixin.com/articles/2019-03-05-8
[2] https://www.cnblogs.com/marsggbo/p/11461155.html
[3] https://baike.baidu.com/item/%E6%AD%A3%E5%AE%9A%E7%9F%A9%E9%98%B5/11030459?fr=aladdin
[4] https://www.qiujiawei.com/linear-algebra-11/
[5] https://blog.csdn.net/ACdreamers/article/details/44656847
[6] 使用GPU加速计算矩阵的Cholesky分解，沈 聪，高火涛
[7] https://blog.csdn.net/zb1165048017/article/details/70207812