大规模线性方程组解法简介

原创 www.cae-sim.com 多物理场仿真技术

前面介绍过主要的线性方程组求解库，参考附录。求解大规模线性方程组是仿真软件求解器的底层技术，求解器时间基本都消耗在方程组求解上。线性方程组的解法比较成熟，方法也有很多，而且不同的方法对应不同类型方程组，所以在方法选择上实际很讲究。

商业软件通常将方法封装起来，用户包括开发人员都接触不到线性方程组求解方法。商业软件内部一般会根据求解规模，求解类型等选择合适的线性方程组求解方法。

有些商业软件开放了部分接口供用户选择；开源软件比如OpenFOAM，以及使用开源软件的平台Simscale等提供了许多选项供用户选择。

本文简单介绍下线性方程组的常用解法。通常将线性方程组表示为：

A*x=b

A为已知N*N的矩阵，通常称为刚度矩阵（刚度是力学中的概念，电磁，热等也习惯性这么称呼），b为已知向量，x为待求向量。解线性方程组的操作基本围绕矩阵A展开。

首先介绍一些相关术语：

1.矩阵条件数

条件数是一个表征矩阵稳定特性的标志，条件数越大，说明矩阵越不稳定，即当矩阵中数据出现微小变化时，x结果变化非常大。Matlab中可直接使用命令cond(A)查看矩阵条件数。

2.满秩矩阵

用初等行变换将矩阵A化为阶梯形矩阵, 矩阵中非零行的个数就定义为这个矩阵的秩，记做r(A)

即如果矩阵A为N*N，r(A)=N，则矩阵A为满秩矩阵

在边界元，矩量法等计算方法中，最终形成的A为满秩矩阵。

3.稀疏矩阵

矩阵中绝大部分元素都是0

4.对称矩阵

矩阵的上三角和下三角关于对角线对称

5.求解线性方程组直接法

先求出矩阵A的逆矩阵，再乘以向量b

6.求解线性方程组迭代法

简单讲，就是给出一个初始解x'，带入原方程中，每次评价差距逐步修正，直到最终A*x'-b 接近0为止。

7. OOC（out of core）

对于大型方程组，内存通常无法装下，在求解过程中需要将部分数据写到硬盘上，再次使用的时候再读回来。

常用直接法：

直接法的本质是要计算出A的逆矩阵，通常在求解小规模和特征值问题是可以考虑使用直接法。

1. 高斯消去法/Doolittle /三角分解法/追赶法

这是最基本的方法，时间复杂度和空间复杂度都是N的三次方，软件一般都不会使用。

2. 矩阵分解相关

2. 1.LU分解法

LU分解就是将矩阵分解成单位下三角矩阵L和上三角矩阵U，本质上仍然属于高斯消去法

2.2. Cholesky分解

Cholesky 分解是把一个对称正定的矩阵表示成一个下三角矩阵L和其转置的乘积的分解。如果矩阵是正定的，使用 Cholesky分解会比LU分解更加高效。

2.3. LDLT分解法

Cholesky 分解法的改进

2.4.QR分解

QR分解是把矩阵分解成一个正交矩阵与一个上三角矩阵的积。

2.5. Schur分解

2.6. SVD/GSVD

奇异值分解/一般奇异值分解

在求解大规模线性方程组中，一般不会使用直接法求解，但在使用迭代法过程中需要使用直接法中的方法加工数据。

间接法的本质是迭代，不同方法的区别在于如何选取初始值以及迭代方法。

1. 牛顿迭代，Jacobi(雅克比)迭代，Gauss-Seidel迭代

线性迭代方法，一般情况下收敛太慢，大规模方程组不推荐使用。

2. JOR -- Jacobi Over Relaxation

Jacobi 方法加入松弛因子

3. Lanczos方法

适用于稀疏矩阵特征值问题

4. 共轭梯度方法以及改进方法

共轭梯度法

Conjugate gradient method--CG

双共轭梯度法

Bi-conjugate gradient method--BCG

稳定双共轭梯度法

Bi-conjugate gradient stabilized method--GCGS

预条件共轭梯度法

Preconditioned gradient stabilized method--PCG

相比牛顿和Jacobi，通过优化使其共轭的求解向量和方向，加速了求解性能。

共轭梯度法的思想就是找到N个两两共轭的方向，每次沿着一个方向优化得到该方向上的极小值，后面再沿其它方向求极小值的时候，不会影响前面已经得到的沿哪些方向上的极小值，所以理论上对n个方向都求出极小值就得到了N维问题的极小值。

5. GMRAS广义最小残量

(Generalized Minimal Residual Algorithm)

非对称系统的线性方程组的数值解迭代法，该方法与最小残量的 Krylov 子空间中向量来逼近解，是求解大规模线性方程组的常用算法之一,具有收敛速度快、稳定性好等优点。

改进的GMRAS：

SGMRAS Simpler GMRAS

PGMRAS Preconditioned GMRAS

6. 快速多级（Fast Multi pole Method）

当矩阵为满秩矩阵时，传统计算方法资源需求和求解规模呈指数级上升，大规模的系统求解需要使用快速多级方法。本技术博客和公众号中有详细介绍。

改进方法：多层快速多级

7. Successive Over Relaxation(SOR)连续松弛

连续过度松弛（SOR）方法是高斯-赛德尔（Gauss-Seidel）方法的一种变体，用于求解线性方程组，从而可以更快地收敛。任何缓慢收敛的迭代过程都可以使用类似的方法。

改进的SOR：

Accelerated Over relaxation (AOR)过松弛

Preconditioned AOR -- PAOR 预条件过松弛

Quasi AOR -- QAOR准过松弛

8. Krylov子空间迭代法

将矩阵A分解成多个子矩阵，并用一系列线性表达式组合表示。将整个系统降维，利于并行计算，是大规模线性方程组有效的一种解法。其中涉及到了Lanczos方法和Arnoldi方法。

按照矩阵A的特点，我们可以做如下分类：

1. 是否是满秩矩阵；

2. 是否是稀疏矩阵；

3. 是否是对称矩阵；

4. 是否是病态矩阵；

5. 是否是正定矩阵

6. 矩阵规模（矩阵中N的大小）

在选择求解方法的时候，需要考虑到方程组矩阵以上特点。根据作者的经验，方程组的求解性能高度依赖矩阵特征，矩阵规模和硬件。