Introduction to Mathematical Programming

Ming Zhong

Lecture 7

September 14, 2018

Table of Contents

Solving Linear Systems as Finding A Fixed Point

We want to solve the linear system, $\mathbf{A}\vec{x} = \vec{y}$, when $\mathbf{A}$ is difficult to invert directly, so we consider

$$\begin{aligned}\mathbf{A}\vec{x} &= \vec{y} \\ \vec{\mathbf{0}} &= \vec{y} - \mathbf{A}\vec{x} \\ \mathbf{B}\vec{x} &= \vec{y} + (\mathbf{B} - \mathbf{A})\vec{x} \\ \vec{x} &= \mathbf{B}^{-1}\vec{y} + \mathbf{B}^{-1}(\mathbf{B} - \mathbf{A})\vec{x}\end{aligned}$$

where the matrix $\mathbf{B}$ is very easy to invert, then we have changed solving a linear system into finding a fixed point.

- The options of choosing different $\mathbf{B}$'s give rise to different iterative solution methods: Jacobi, Gauss-Seidel.
- Other iterative solution methods go under the name of Krylov subspace methods: Conjugate Gradient.
- The idea is: to start with an initial guess and use an iterative procedure that will converge to the solution.

Iterative Methods: Jacobi

The simplest one is to pick $\mathbf{B} = \text{diag}(a_{11}, \cdots, a_{NN})$ with a_{ii}'s being the diagonal entries of $\mathbf{A} \in \mathbb{R}^{N\times N}$. The Jaboci algorithm:

- Make an initial guess, $\vec{x}_0$.
- Apply the Jacobi scheme, $\vec{x}_{k+1} = \mathbf{B}^{-1}(\vec{y} + (\mathbf{B} - \mathbf{A})\vec{x}_k)$.
- Check for convergence: 1) $\left|\left|\vec{x}_{k+1} - \vec{x}_k\right|\right| < \text{tol}$; 2)$\left|\left|\vec{y} - \mathbf{A}\vec{x}_{k+1}\right|\right| < \text{tol}$.

Example

Let us consider the following system,

$$
\begin{aligned}
4x - y + z &= 7 \\
4x - 8y + z &= -21 \\
-2x + y + 5z &= 15.
\end{aligned}
$$

Example, cont.

Example

Re-arrange the terms,

$$x = \frac{7 + y - z}{4}$$
$$y = \frac{21 + 4x + z}{8}$$
$$z = \frac{15 + 2x - y}{5}.$$

then we obtain an iterative scheme,

$$x_{k+1} = \frac{7 + y_k - z_k}{4}$$
$$y_{k+1} = \frac{21 + 4x_k + z_k}{8}$$
$$z_{k+1} = \frac{15 + 2x_k - y_k}{5}.$$

Convergence

Demo.

Another Example

Example

Consider another linear system

$$
\begin{aligned}
-2x + y + 5z &= 15 \\
4x - 8y + z &= -21 \\
4x - y + z &= 7.
\end{aligned}
$$

then we obtain an iterative scheme,

$$
\begin{aligned}
x_{k+1} &= \frac{y_k + 5z_k - 15}{2} \\
y_{k+1} &= \frac{21 + 4x_k + z_k}{8} \\
z_{k+1} &= y_k - 4x_k + 7.
\end{aligned}
$$

Convergence

Demo.

Strictly Diagonally Dominant

Two different systems, two kinds of convergence. what happened?

Definition

A matrix $\mathbf{A} \in \mathbb{R}^{N \times N}$ is called strictly diagonally dominant if the following holds,

$$\left|a_{kk}\right| > \sum_{j=1, j \neq k}^{N} \left|a_{kj}\right|,$$

for $k = 1, \cdots, N$.

Apparently,

$$A = \begin{pmatrix} 4 & -1 & 1 \\ 4 & -8 & 1 \\ -2 & 1 & 5 \end{pmatrix} \text{ is strictly diagonally dominant.}$$

Strictly Diagonally Dominant, cont.

And

$$A = \begin{pmatrix} -2 & -1 & 5 \\ 4 & -8 & 1 \\ 4 & -1 & 5 \end{pmatrix} \text{ is not.}$$

That is why Jacobi converges for the first system, not diverges for the second system.

- Strictly Diagonally Dominant matrices are non-singular (invertible, has an inverse, $\det(\mathbf{A}) \neq 0$).
- non-singular matrices might not be strictly diagonally dominant.

Iterative Methods: Gauss-Seidel

Now if we consider $\mathbf{B} = \mathbf{D} - \mathbf{L}$ where

- $\mathbf{D}$ is a diagonal matrix containing the diagonal entries of $\mathbf{A}$.
- $\mathbf{L}$ is a strict lower triangular matrix $l_{kk} = 0$ containing the negative of the entries of the lower triangular part of $\mathbf{A}$.

Then we obtain the Gauss-Seidel method,

- It converges when $\mathbf{A}$ is strict diagonally dominant.
- It converges a "bit" faster than Jacobi.

Example

Example

Consider another linear system

$$
\begin{aligned}
4x - y + z &= 7 \\
4x - 8y + z &= -21 \\
-2x + y + 5z &= 15.
\end{aligned}
$$

then we obtain an iterative scheme,

$$
\begin{aligned}
x_{k+1} &= \frac{7 + y_k - z_k}{4} \\
y_{k+1} &= \frac{21 + 4x_{k+1} + z_k}{8} \\
z_{k+1} &= \frac{15 + 2x_k - y_{k+1}}{5}.
\end{aligned}
$$

Convergence

Demo.

Other Iterative Methods

Consider $\mathbf{A} = \mathbf{D} - \mathbf{L} - \mathbf{U}$, where

- The diagonal matrix $\mathbf{D}$ contains the diagonal entries of $\mathbf{A}$.
- The strict lower triangular matrix $\mathbf{L}$ contains the negative of the lower triangular part of $\mathbf{A}$.
- The strict upper triangular matrix $\mathbf{U}$ contains the negative of the upper triangular part of $\mathbf{A}$.

Choosing different $\mathbf{B}$ matrix:

- Richardson method: $\mathbf{B} = \frac{1}{\omega} I$ $(\omega \neq 0)$.
- Jacobi method: $\mathbf{B} = \mathbf{D}$.
- Damped Jacobi method: $\mathbf{B} = \frac{1}{\omega}\mathbf{D}$.
- Gauss-Seidel method: $\mathbf{B} = \mathbf{D} - \mathbf{L}$.
- Successive over-relaxation method (SOR): $\mathbf{B} = \frac{1}{\omega}\mathbf{D} - \mathbf{L}$ $(\omega \neq 0)$.
- Symmetric successive over-relaxation (SSOR): $\mathbf{D} = \frac{1}{\omega(\omega-2)}(\mathbf{D} - \omega\mathbf{L})\mathbf{D}^{-1}(\mathbf{D} - \omega\mathbf{U})$ $(\omega \neq 0, 2)$.

Krylov Subspace Method

Definition

The order r Krylov subspace generated by an $N \times N$ matrix $\mathbf{A}$ and a vector $\vec{y}$ of dimension N is the linear subspace spanned by the images of $\vec{y}$ under the first r powers of $\mathbf{A}$, that is,

$$\mathcal{K}_r(\mathbf{A}, \vec{y}) = \text{span}\{\vec{y}, \mathbf{A}\vec{y}, \mathbf{A}^2\vec{y}, \cdots, \mathbf{A}^{r-1}\vec{y}\}$$

Krylove subspace methods work,

- by forming a basis of the sequence of successive matrix powers times the initial residual (the Krylov sequence).
- the approximation of the solution are then formed by minimizing the residual over the subspace formed.
- The prototypical method is the conjugate gradient (CG) method (assuming $\mathbf{A}$ is symmetric positive-definite).

Krylov Subspace Method, cont.

Continue on,

- For symmetric (possibly indefinite) $\mathbf{A}$, we use the minimal residual method (MINRES).
- Not symmetric $\mathbf{A}$, we use generalized minimal residual method (GMRES) and biconjugate gradient method (BiCG).

Convergence,

- Since we are forming a basis, it stops after N iterations (in exact arithmetic).
- with round off errors, it might not.
- When N is large, the approximation is usually "good" after just a few iterations.
- the procedure can be sped up by using pre-conditioner.

MATLAB Routines

Krylov subspace methods implemented in MATLAB,

- Conjugate Gradient method (CG): "cg".
- Minimal Residual method (MINRES): "minres".
- Generalized Minimal Residual method (GMRES): "gmres".
- Bi-conjugate gradient method (BiCG): "bicg".