Numerical Analysis

Mathematics of Scientific Computing

主讲人 邱欣欣
幻灯片制作 邱欣欣

中国海洋大学 信息科学与工程学院

2013 年 11 月 10 日

Solving Systems of Linear Equations

Contents

Introduction

Solving the system

$$Ax = b$$

for the case when A is a real $n \times n$, symmetric, and positive definite matrix. These means that

$$A^T = A$$

And

$$x^T Ax > 0 \quad \text{for} \quad x \neq 0$$

Lemma on Quadratic Form

LEMMA 1 (Lemma on Quadratic Form)

If A is symmetric and positive definite, then the problem of solving $Ax = b$ is equivalent to the problem of minimizing the quadratic form

$$q(x) = \langle x, Ax \rangle - 2\langle x, b \rangle$$

- Soving $Ax = b \Longrightarrow$ the minimum of quadratic functional $\Longrightarrow$ $x^{(k+1)} = x^{(k)} + t_k v^{(k)}$

Introduction

Proof.

$$
\begin{aligned}
q(x) &= \langle x, Ax\rangle - 2\langle x, b\rangle \\
&= x^T A x - 2 b^T x
\end{aligned}
$$

Computing the derivative

$$
\frac{\partial q}{\partial x_i} = 2(a_{i1}x_1 + \ldots + a_{in}x_n) - 2b_i \qquad i = 1, 2, \ldots, n
$$

$$
\text{grad } q(x) = 2(Ax - b) = -2r \qquad r = b - Ax
$$

$\square$

Introduction

$$\text{grad } q(x) = 2(Ax - b) = -2r \qquad r = b - Ax$$

- If $q(x)$ reaches its minimum at x^*, then grad $q(x) = 0$, so $Ax^* = b$ and x^* is the solution of system.
- If x^* is a solution of the system $Ax = b$, for any vector y

$$\begin{aligned} q(x^* + y) &= (x^* + y)^T A(x^* + y) - 2b^T(x^* + y) \\ &= {x^*}^T Ax^* - 2b^T x^* + y^T Ay = q(x^*) + y^T Ay \end{aligned}$$

When A is positive definite matrix, $y^T Ay \geqslant 0$, so $q(x^* + y) \geqslant q(x^*)$, x^* is a minimum point of q.

Introduction

- $$\begin{aligned} x &= x^0 + tv^0 \\ x^1 &= x^0 + t_0 v^0 \\ q(x^0 + t_0 v^0) &\leqslant q(x^0 + tv^0) \end{aligned}$$

$$\begin{aligned} x &= x^k + tv^k \\ q(x^k + t_k v^k) &\leqslant q(x^k + tv^k) \\ x^{k+1} &= x^k + t_k v^k \end{aligned}$$

Introduction

- Now we should compute t.

$$\begin{aligned} q(x+tv) &= \langle x+tv, A(x+tv)\rangle - 2\langle x+tv, b\rangle \\ &= q(x) + 2t\langle v, Ax-b\rangle + t^2\langle v, Av\rangle \end{aligned}$$

Computing the derivative

$$\frac{d}{dt}q(x+tv) = 2\langle v, Ax-b\rangle + 2t\langle v, Av\rangle$$

So the value of t that yields the minimum point is

$$\hat{t} = \langle v, b-Ax\rangle / \langle v, Av\rangle$$

Introduction

- Then we obtained the general form of iterative methods

$$x^{(k+1)} = x^{(k)} + t_k v^{(k)}$$

where

$$t_k = \frac{\langle v^{(k)}, b - Ax^{(k)} \rangle}{\langle v^{(k)}, Av^{(k)} \rangle}$$

$$r^{(k)} = b - Ax^{(k)}$$

Contents

Steepest Descent

- In the method of steepest descent, $v^{(k)}$ should be the negative gradient of q at $x^{(k)}$. And the negative gradient points in the direction of the residual, $r^{(k)} = b - Ax^{(k)}$.
- input x, A, b, M
  output $0, x$
  for $k = 1$ to M do
  $v \leftarrow b - Ax$
  $t \leftarrow \langle v, v\rangle / \langle v, Av\rangle$
  $x \leftarrow x + tv$
  output k, x
  end do

Steepest Descent

```
function x=SteepestDescent(x,A,b,M)
for k=1:M
    v=b-A*x;
    t=dot(v,v)/dot(v,A*v);
    x=x+t*v;
end
```

Contents

Conjugate Gradient Method

- Given initial vector x^0, at step 1 we choose the negative gradient of q at x^0 as the v^0, and $v^0 = r^0$.
- At $k+1$th ($k \geqslant 1$) step, we choose v_k from two-dimensional plane formed by r_k and v_{k-1}

$$\pi_2 = \{x = x_k + \xi r_k + \eta v_{k-1} : \xi, \eta \in \mathbb{R}\}$$

Considering the limitation on plane π_2

$$\begin{aligned} \psi(\xi, \eta) &= q(x_k + \xi r_k + \eta v_{k-1}) \\ &= (x_k + \xi r_k + \eta v_{k-1})^T A (x_k + \xi r_k + \eta v_{k-1}) \\ &\quad -2b^T (x_k + \xi r_k + \eta v_{k-1}) \end{aligned}$$

Conjugate Gradient Method

- Let $\frac{\partial \psi}{\partial \xi} = \frac{\partial \psi}{\partial \eta} = 0$. Then q has a unique minimum point in π_2

$$\tilde{x} = x_k + \xi_0 r_k + \eta_0 v_{k-1}$$

The ξ_0 and η_0 satisfy

$$\begin{cases} \xi_0 r_k^T A r_k + \eta_0 r_k^T A v_{k-1} = r_k^T r_k \\ \xi_0 r_k^T A v_{k-1} + \eta_0 v_{k-1}^T A v_{k-1} = 0 \end{cases}$$

we can select $v_k = \frac{1}{\xi_0}(\tilde{x} - x_k) = r_k + \frac{\eta_0}{\xi_0} v_{k-1}$ as the new search direction.

Let $s_{k-1} = \frac{\eta_0}{\xi_0} = -\frac{r_k^T A v_{k-1}}{v_{k-1}^T A v_{k-1}}$

Conjugate Gradient Method

- Notice that v_k satisfy $v_k^T A v_{k-1} = 0$, it means v_k and v_{k-1} mutually conjugate.
-

$$
\begin{aligned}
t_k &= \frac{r_k^T v_k}{v_k^T A v_k} = \frac{r_k^T r_k}{v_k^T A v_k} \\
s_k &= -\frac{r_{k+1}^T A v_k}{v_k^T A v_k} = \frac{r_{k+1}^T r_{k+1}}{r_k^T r_k}
\end{aligned}
$$

$$
\begin{aligned}
x_{k+1} &= x_k + t_k v_k \\
r_{k+1} &= b - A x_{k+1} = r_k - t_k A v_k \\
v_{k+1} &= r_{k+1} + s_k v_k
\end{aligned}
$$

We will use $r_k^T r_{k+1} = r_k^T v_{k-1} = r_{k+1}^T v_k = 0,\ k = 1, 2, \ldots$.

Conjugate Gradient Method

- input $x^{(0)}, M, A, b, \varepsilon$
  $r^{(0)} \leftarrow b - Ax^{(0)}$
  $v^{(0)} \leftarrow r^{(0)}$
  output $0, x^{(0)}, r^{(0)}$
  for $k = 0$ to $M - 1$ do
      if $v^{(k)} = 0$ then stop
      $t_k \leftarrow \langle r^{(k)}, r^{(k)}\rangle / \langle v^{(k)}, Av^{(k)}\rangle$
      $x^{(k+1)} \leftarrow x^{(k)} + t_k v^{(k)}$
      $r^{(k+1)} \leftarrow r^{(k)} - t_k Av(k)$
      if $\| r^{(k+1)} \|_2^2 < \varepsilon$ then stop
      $s_k \leftarrow \langle r^{(k+1)}, r^{(k+1)}\rangle / \langle r^{(k)}, r^{(k)}\rangle$
      $v^{(k+1)} \leftarrow r^{(k+1)} + s_k v^{(k)}$
      output $k + 1, x^{(k+1)}, r^{(k+1)}$
  end do

- input $x, A, b, M, Q, \delta, \varepsilon$
  $r \leftarrow b - Ax$
  $v \leftarrow r$
  $c \leftarrow \langle r, r \rangle$
  for $k = 1$ to M do
   if $\langle v, v \rangle^{1/2} < \delta$ then exit loop
   $z \leftarrow Av$
   $t \leftarrow c/\langle v, z \rangle$
   $x \leftarrow x + tv$
   $r \leftarrow r - tz$
   $d \leftarrow \langle r, r \rangle$
   if $d < \varepsilon$ then exit loop
   $v \leftarrow r + (d/c)v$
   $c \leftarrow d$
   output k, x, r
  end do

```
function x=ConGradient(x,A,b,M,epsilon,△)
r=b-A*x;
v=r;
c=dot(r,r);
for k=1:M
    if sqrt(dot(v,v))<△
        break
    end
    z=A*v;
    t=c/dot(v,z);
    x=x+t*v;
    r=r-t*z;
    d=dot(r,r)
    if d<epsilon
        break
    end
    v=r+(d/c)*v;
    c=d;
end
```

Conjugate Gradient Method

- The **conjugate gradient** method, the search directions $v^{(i)}$ form an A-orthogonal system, that is, $\langle v^{(i)}, Av^{(j)}\rangle = 0$ if $i \neq j$.
- The property is that the residuals, $r^{(i)} = b - Ax^{(i)}$, form an orthogonal system, that is, $\langle r^{(i)}, r^{(j)}\rangle = 0$ if $i \neq j$.
- Theoretically, the conjugate gradient algorithm will yield the solution of system $Ax = b$ in at most n steps.

Conjugate Gradient Method

THEOREM 1 (Theorem on Conjugate Gradient Algorithm)

In the conjugate gradient algorithm, for any integer $m < n$, if $v^{(0)}, v^{(1)}, \ldots, v^{(m)}$ are all nonzero vectors, then $r^{(i)} = b - Ax^{(i)}$ for $0 \leqslant i \leqslant m$, and $\{r^{(0)}, r^{(1)}, \ldots, r^{(m)}\}$ is an orthogonal set of nonzero vectors.

Proof.

1. $\langle r^{(m)}, v^{(i)} \rangle = 0 \qquad (0 \leqslant i < m)$
2. $\langle r^{(i)}, r^{(i)} \rangle = \langle r^{(i)}, v^{(i)} \rangle \qquad (0 \leqslant i \leqslant m)$
3. $\langle v^{(m)}, Av^{(i)} \rangle = 0 \qquad (0 \leqslant i < m)$
4. $r^{(i)} = b - Ax^{(i)} \qquad (0 \leqslant i \leqslant m)$
5. $\langle r^{(m)}, r^{(i)} \rangle = 0 \qquad (0 \leqslant i < m)$
6. $r^{(i)} \neq 0 \qquad (0 \leqslant i \leqslant m)$

Proof.

The proof is by induction on m.

For the case $m = 0$, we assume that $v^{(0)} \neq 0$, then $r^{(0)} = b - Ax^{(0)} = v^{(0)} \neq 0$.

Assume that the theorem is true for m, we shall prove it for $m + 1$.

1. $\langle r^{(m+1)}, v^{(i)} \rangle = 0 \qquad (0 \leqslant i \leqslant m)$.

Let $i = m$,

$$\begin{aligned}
\langle r^{(m+1)}, v^{(m)} \rangle &= \langle r^{(m)} - t_m A v^{(m)}, v^{(m)} \rangle \\
&= \langle r^{(m)}, v^{(m)} \rangle - t_m \langle v^{(m)}, A v^{(m)} \rangle \\
&= \langle r^{(m)}, v^{(m)} \rangle - \langle r^{(m)}, r^{(m)} \rangle = 0
\end{aligned}$$

If $0 \leqslant i < m$,

$$\langle r^{(m+1)}, v^{(i)} \rangle = \langle r^{(m)}, v^{(i)} \rangle - t_m \langle v^{(m)}, A v^{(i)} \rangle = 0$$

Conjugate Gradient Method

Proof.

2. $\langle r^{(m+1)}, r^{(m+1)} \rangle = \langle r^{(m+1)}, v^{(m+1)} \rangle$.

$$\langle r^{(m+1)}, v^{(m+1)} \rangle = \langle r^{(m+1)}, r^{(m+1)} + s_m v^{(m)} \rangle = \langle r^{(m+1)}, r^{(m+1)} \rangle$$

□

Contents

Preconditioned Conjugate Gradient

- For $Ax = b$, preconditioning this system and obtain a new system that is better conditioned. So for some nonsingular matrix S, the preconditioned system
$$\widehat{A}\hat{x} = \hat{b}$$
where
$$\begin{cases} \widehat{A} = S^T A S \\ \hat{x} = S^{-1} x \\ \hat{b} = S^T b \end{cases}$$
- Suppose that the symmetric and positive definite splitting matrix Q can be factored so that
$$Q^{-1} = SS^T$$

Preconditioned Conjugate Gradient

- We write

$$\begin{aligned}
\hat{x}^{(k)} &= S^{-1}x^{(k)} \\
\hat{v}^{(k)} &= S^{-1}v^{(k)} \\
\hat{r}^{(k)} &= \hat{b} - \widehat{A}\hat{x}^{(k)} = S^T b - (S^T A S)S^{-1}x^{(k)} = S^T r^{(k)} \\
\tilde{r}^{(k)} &= Q^{-1}r^{(k)}
\end{aligned}$$

Then, we obtain

$$\begin{aligned}
\hat{t}_k &= \langle \hat{r}^{(k)}, \hat{r}^{(k)} \rangle / \langle \hat{v}^{(k)}, \widehat{A}\hat{v}^{(k)} \rangle = \langle \tilde{r}^{(k)}, r^{(k)} \rangle / \langle v^{(k)}, Av^{(k)} \rangle \\
x^{(k+1)} &= x^{(k)} + \hat{t}_k v^{(k)} \\
r^{(k+1)} &= r^{(k)} - \hat{t}_k A v^{(k)} \\
\hat{s}_k &= \langle \hat{r}^{(k+1)}, \hat{r}^{(k+1)} \rangle / \langle \hat{r}^{(k)}, \hat{r}^{(k)} \rangle = \langle \tilde{r}^{(k+1)}, r^{(k+1)} \rangle / \langle \tilde{r}^{(k)}, r^{(k)} \rangle \\
v^{(k+1)} &= \tilde{r}^{(k+1)} + \hat{s}_k v^{(k)}
\end{aligned}$$

- input $x, A, b, M, Q, \delta, \varepsilon$
  $r \leftarrow b - Ax$
  solve $Qz = r$ for z
  $v \leftarrow z$; $c \leftarrow \langle z, r\rangle$
  for $k = 1$ to M do
   if $\langle v, v\rangle^{1/2} < \delta$ then exit loop
   $z \leftarrow Av$; $t \leftarrow c/\langle v, z\rangle$
   $x \leftarrow x + tv$; $r \leftarrow r - tz$
   solve $Qz = r$ for z
   $d \leftarrow \langle z, r\rangle$
   if $d < \varepsilon$ then
    if $\langle r, r\rangle < \varepsilon$ then exit loop
   end if
   $v \leftarrow z + (d/c)v$
   $c \leftarrow d$
   output k, x, r
  end do

Preconditioned Conjugate Gradient

```
function x=PCG(x,A,b,M,Q,△,epsilon)
r=b-A*x
n=length(Q);
for k=1:n
    z(k)=r(k)/Q(k,k)
end
v=z'
c=dot(z,r)
for k=1:M
    if sqrt(dot(v,v))<△
        break
    end
    z=A*v;
    t=c/dot(v,z);
    x=x+t*v;
    r=r-t*z;
```

Preconditioned Conjugate Gradient

```
for k=1:n
    z(k)=r(k)/Q(k,k);
    end
    d=dot(z,r);
    if d<epsilon
        if dot(r,r)<epsilon
            break
        end
    end
    v=z+(d/c)*v;
    c=d;
end
```

Contents

Analysis of Roundoff Error in the Gaussian Algorithm

THEOREM 2 (Theorem on $\tilde{L}\tilde{U}$-Factorization)

Let A be an $n \times n$ nonsingular matrix whose elements are machine numbers in a computer with unit roundoff ε. The Gaussian algorithm with row pivoting produces matrices $\tilde{L}$ and $\tilde{U}$ such that

$$\tilde{L}\tilde{U} = A + E \quad \textit{where} \quad |e_{ij}| \leqslant 2n\varepsilon \max_{1 \leqslant i,j,k \leqslant n} |a_{ij}^{(k)}|$$

Analysis of Roundoff Error in the Gaussian Algorithm

THEOREM 3 (Theorem on Roundoff Error in Dot Product)

If $x_1, x_2, \ldots, x_n$ and $y_1, y_2, \ldots, y_n$ are machine numbers, then the machine value of

$$\sum_{i=1}^{n} x_i y_i$$

computed in the natural way can be expressed as $\sum_{i=1}^{n} x_i y_i (1+\delta_i)$, in which the δ_i's satisfy $|\delta_i| \leqslant \frac{6}{5}(n+1)\varepsilon$. (The number ε is the unit roundoff error of the machine, and we assume that $n\varepsilon < \frac{1}{3}$.)

Analysis of Roundoff Error in the Gaussian Algorithm

THEOREM 4 (Theorem on Perturbed Unit Lower Triangular System)

Let L be an $n \times n$ unit lower triangular matrix whose elements are machine numbers. Let b be a vector whose components are machine numbers. The computed solution of $Ly = b$ is a vector $\tilde{y}$ that is the exact solution of

$$(L + \Delta)\tilde{y} = b \quad \textit{with} \quad |\Delta_{ij}| \leqslant \frac{6}{5}(n+1)\varepsilon|l_{ij}|$$

Here ε is the machine's unit roundoff error, and it is assumed that $n\varepsilon < \frac{1}{3}$.

Analysis of Roundoff Error in the Gaussian Algorithm

THEOREM 5 (Theorem on Perturbed Upper Triangular System)

Let U be an $n \times n$, upper triangular, nonsingular matrix. If the elements of U and c are machine numbers, and if $n\varepsilon < \frac{1}{3}$, then the computed solution $\tilde{y}$ of $Uy = c$ satisfies exactly a perturbed system

$$(U + \Delta)\tilde{y} = c \quad \textit{with} \quad |\Delta_{ij}| \leqslant \frac{6}{5}(n+1)\varepsilon|u_{ij}|$$

Analysis of Roundoff Error in the Gaussian Algorithm

THEOREM 6 (Theorem on Perturbed System)

Let the elements of A and b be machine numbers. If the Gaussian algorithm with row pivoting is used to solve $Ax = b$, then the computed solution $\tilde{x}$ is the exact solution of a perturbed system

$$(A+F)\tilde{x} = b \quad \text{in which} \quad |f_{ij}| \leqslant 10n^2\varepsilon\rho$$

Here n is the order of the matrix A, $\rho = \max_{1\leqslant i,j,k\leqslant n} |a_{ij}^{(k)}|$, and ε is the machine's unit roundoff error. It is assumed that $n\varepsilon < \dfrac{1}{3}$.