Math 1080: Numerical Linear Algebra
Chapter 5, Solving $Ax = b$ by Optimization

M. M. Sussman

sussmanm@math.pitt.edu
Office Hours: MW 1:45PM-2:45PM, Thack 622

March 2015
Some definitions

- A matrix $A$ is "SPD" (Symmetric Postive Definite)
  1. $A = A^T$
  2. $x \cdot Ax = \langle x, Ax \rangle = x^T Ax > 0$ for $x \neq 0$
- $A$ is “negative definite” if $-A$ is SPD
- $A$ is “skew-symmetric” if $A^T = -A$
- $A$ is “indefinite” if symmetric and neither SPD nor negative definite.
- Two symmetric matrices satisfy $A \succ B$ if $(A - B)$ is SPD
Important facts

- All eigenvalues of a symmetric matrix are real
- A symmetric matrix has an orthogonal basis of eigenvectors
- A symmetric matrix $A$ is positive definite if and only if all $\lambda(A) > 0$. 
If $A$ is SPD then

$$\langle x, y \rangle_A = x^t Ay$$

is a weighted inner product on $\mathbb{R}^n$, the $A$-inner product, and

$$\|x\|_A = \sqrt{\langle x, x \rangle_A} = \sqrt{x^t Ax}$$

is a weighted norm, the $A$-norm.
Proof

bilinear

\[ \langle u+v, y \rangle_A = (u+v)^t Ay = u^t Ay + v^t Ay = \langle u, y \rangle_A + \langle v, y \rangle_A \]

and

\[ \langle \alpha u, y \rangle_A = (\alpha u)^t Ay = \alpha (u^t Ay) = \alpha \langle u, y \rangle_A. \]

Positive

\[ \langle x, x \rangle_A = x^t Ax > 0, \text{ for } x \neq 0, \text{ since } A \text{ is SPD}. \]

Symmetric

\[ \langle x, y \rangle_A = x^t Ay = (x^t Ay)^t = y^t A^t x^t = y^t Ax = \langle y, x \rangle_A. \]

Hence, \( \langle x, x \rangle_A \) induces a norm \( \|x\|_A = \sqrt{\langle x, x \rangle_A} \)
The $A$-norm is a weighted 2-norm

The norm induced by a $2 \times 2$ SPD matrix $A$ is just a weighted 2-norm.

Proof:

- Let $(\lambda_1, \phi_1)$ and $(\lambda_2, \phi_2)$ be the two eigenvalues and eigenvectors of $A$.
- Any vector $v = \alpha_1 \phi_1 + \alpha_2 \phi_2$
- $\|v\|_A^2 = (\alpha_1 \phi_1 + \alpha_2 \phi_2)^T A (\alpha_1 \phi_1 + \alpha_2 \phi_2)$
- $\|v\|_A^2 = (\alpha_1 \phi_1 + \alpha_2 \phi_2)^T (\alpha_1 \lambda_1 \phi_1 + \alpha_2 \lambda_2 \phi_2)$
- $\|v\|_A^2 = (\alpha_1^2 \lambda_1 + \alpha_2^2 \lambda_2)$
Homework

Text, Exercises 258, 259
The connection to optimization

Descent Methods

Application to Stationary Iterative Methods

Parameter selection

The Steepest Descent Method
Finding a solution = finding a minimum

For A SPD, the unique solution of

\[ Ax = b \]

is also the unique minimizer of the functional

\[ J(x) = \frac{1}{2} \langle x, Ax \rangle - \langle x, b \rangle = \frac{1}{2} (x^T Ax) - x^T b \]

**Proof**

Take derivatives of both sides.

\[ \frac{\partial J}{\partial x_i} = 0 = \frac{1}{2} ((Ax)_i + (Ax)_i) - b_i \]
$x = A^{-1}b$ is the argument that minimizes $J(y)$:

$$x = \arg \min_{y \in \mathbb{R}^N} J(y)$$
Theorem 260

Let $A$ be SPD. The solution of $Ax = b$ is the unique minimizer of $J(x)$.

- For any $y$
  \[ J(x + y) = J(x) + \frac{1}{2} y^t Ay > J(x) \]

- And
  \[ \|x - y\|_A^2 = 2 (J(y) - J(x)) . \]
Proof of first claim

\[ J(x + y) = \frac{1}{2} (x + y)^t A (x + y) - (x + y)^t b \]

\[ = \frac{1}{2} x^t Ax + y^t Ax + \frac{1}{2} y^t Ay - x^t b - y^t b \]
Proof of first claim

\[ J(x + y) = \frac{1}{2} (x + y)^t A (x + y) - (x + y)^t b \]

\[ = \frac{1}{2} x^t A x + y^t A x + \frac{1}{2} y^t A y - x^t b - y^t b \]

\[ = \left( \frac{1}{2} x^t A x - x^t b \right) + \left( y^t A x - y^t b \right) + \frac{1}{2} y^t A y \]
Proof of first claim

\[ J(x + y) = \frac{1}{2}(x + y)^t A(x + y) - (x + y)^t b \]

\[ = \frac{1}{2}x^t Ax + y^t Ax + \frac{1}{2}y^t Ay - x^t b - y^t b \]

\[ = (\frac{1}{2}x^t Ax - x^t b) + (y^t Ax - y^t b) + \frac{1}{2}y^t Ay \]

\[ = J(x) + y^t (Ax - b) + \frac{1}{2}y^t Ay \]
Proof of first claim

\[ J(x + y) = \frac{1}{2}(x + y)^t A(x + y) - (x + y)^t b \]
\[ = \frac{1}{2}x^t Ax + y^t Ax + \frac{1}{2}y^t Ay - x^t b - y^t b \]
\[ = (\frac{1}{2}x^t Ax - x^t b) + (y^t Ax - y^t b) + \frac{1}{2}y^t Ay \]
\[ = J(x) + y^t (Ax - b) + \frac{1}{2}y^t Ay \]
\[ = J(x) + \frac{1}{2}y^t Ay > J(x). \]
Proof of first claim

\[ J(x + y) = \frac{1}{2} (x + y)^t A (x + y) - (x + y)^t b \]

\[ = \frac{1}{2} x^t Ax + y^t Ax + \frac{1}{2} y^t Ay - x^t b - y^t b \]

\[ = \left( \frac{1}{2} x^t Ax - x^t b \right) + \left( y^t Ax - y^t b \right) + \frac{1}{2} y^t Ay \]

\[ = J(x) + y^t (Ax - b) + \frac{1}{2} y^t Ay \]

\[ = J(x) + \frac{1}{2} y^t Ay > J(x). \]

\[ A \text{ is SPD, so } y \neq 0 \implies y^t Ay > 0. \]
Proof of second claim

LHS

\[ \|x - y\|_A^2 \]
\[ = (x - y)^t A(x - y) \]
\[ = x^t Ax - y^t Ax - x^t Ay + y^t Ay \]
\[ = (\text{since } Ax = b) \]
\[ = x^t b - y^t b - y^t b + y^t Ay \]
\[ = y^t Ay - 2y^t b + x^t b \]

RHS

\[ 2 (J(y) - J(x)) \]
\[ = y^t Ay - 2y^t b - x^t Ax + 2x^t b \]
\[ = (\text{since } Ax = b) \]
\[ = y^t Ay - 2y^t b - x^t [Ax - b] + x^t b \]
\[ = y^t Ay - 2y^t b + x^t b \]
The $2 \times 2$ case

- $2 \times 2$ system $Ax = b$ with

$$ A = \begin{bmatrix} a & c \\ c & d \end{bmatrix}. $$

- Eigenvalues: $(a - \lambda)(d - \lambda) - c^2 = 0$

$$ \lambda = \frac{(a + d) \pm \sqrt{(a + d)^2 - 4(ad - c^2)}}{2} $$

- Positive definite ($\lambda > 0$) $\implies$ $(a + d) > 0$ and $c^2 - ab < 0$
The 2 × 2 case cont’d

Write \( x = [x_1, x_2]^T \)

\[
J(x_1, x_2) = \frac{1}{2}(x_1, x_2) \begin{bmatrix} a & c \\ c & d \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} - [x_1, x_2] \begin{bmatrix} b_1 \\ b_2 \end{bmatrix}
\]

\[
= \frac{1}{2}(ax_1^2 + 2cx_1x_2 + dx_2^2) - (b_1x_1 + b_2x_2).
\]

The surface \( z = J(x_1, x_2) \) is a paraboloid opening up if and only if \( a > 0, d > 0, \) and \( c^2 - ad < 0. \)

Hessian matrix is SPD (Theorem 260)
Paraboloid opening upward (again)

- Consider \( b_1 = b_2 = 0 \): \( x_1, x_2 \) are components of the error
- \( J(x_1, x_2) = \frac{1}{2}(ax_1^2 + 2cx_1x_2 + dx_2^2) \)
- Set \( x_1 = \frac{r}{\sqrt{a}} \cos \theta \) and \( x_2 = \frac{r}{\sqrt{d}} \sin \theta \)

\[
J = r^2 \cos^2 \theta + r^2 \sin^2 \theta + 2r^2 \frac{c}{\sqrt{ad}} \sin \theta \cos \theta
\]

\[
= r^2 \left( 1 + \frac{c}{\sqrt{ad}} \sin 2\theta \right)
\]

- \( J > 0 \) if and only if \( \pm \frac{c}{\sqrt{ad}} < 1 \) or \( c^2 < ad \).
Topics

The connection to optimization

Descent Methods

Application to Stationary Iterative Methods

Parameter selection

The Steepest Descent Method
General Descent Method, Algorithm 263

Given
1. \( Ax = b \)
2. a quadratic functional \( J(\cdot) \) that is minimized at \( x = A^{-1}b \)
3. a maximum number of iterations \( \text{itmax} \)
4. an initial guess \( x^0 \):

Compute \( r^0 = b - Ax^0 \)

for \( n=1:\text{itmax} \)

(\*) Choose a direction vector \( d^n \)
Find \( \alpha = \alpha_n \) by solving the 1d minimization problem:

(\**\*) \( \alpha_n = \arg \min_\alpha J(x^n + \alpha d^n) \)

\( x^{n+1} = x^n + \alpha_n d^n \)
\( r^{n+1} = b - Ax^{n+1} \)

if converged, exit, end

end
Steepest descent

- Functional: $J(x) = \frac{1}{2} x^t Ax - x^t b$
- Descent direction: $d^n = -\nabla J(x^n)$
\(-\nabla J = r\) when \(J(x) = \frac{1}{2} x^t A x - x^t b\)

- \(J = \frac{1}{2} x^T A x - x^T b\)
- \(\frac{\partial J}{\partial x_i} = \frac{1}{2} (A x)_i + \frac{1}{2} (x^T A)_i - b_i\)
- But \(\frac{1}{2} (x^T A)_i = \frac{1}{2} ((A x)^T)_i = \frac{1}{2} (A x)_i\)
- So \(\frac{\partial J}{\partial x_i} = (A x)_i - b_i = -r_i\)
- \(-\nabla J = r\)
Formula for $\alpha$

Proof of following is a homework problem:
For arbitrary $d^n$,

$$\alpha_n = \frac{(d^n)^T r^n}{(d^n)^T A d^n}$$
Proof of following is a homework problem:
For arbitrary $d^n$,

$$\alpha_n = \frac{(d^n)^T r^n}{(d^n)^T A d^n}$$

Hint: Set $\frac{d}{d\alpha}(J(x + \alpha d)) = 0$ and solve.
Homework

Text, 264.
Use the $\langle \cdot, \cdot \rangle$ notation

$\alpha$ for steepest descent becomes

$$\alpha_n = \frac{\langle d^n, r^n \rangle}{\langle d^n, Ad^n \rangle} = \frac{\langle d^n, r^n \rangle}{\langle d^n, d^n \rangle_A}.$$
Descent demonstration

descentdemo.m
Different descent methods: descent direction $d^n$

- Steepest descent direction: $d^n = -\nabla J(x^n)$
- Random directions: $d^n = \text{a randomly chosen vector}$
- Gauss-Seidel like descent: $d^n$ cycles through the standard basis of unit vectors $e_1, e_2, \cdots, e_N$ and repeats if necessary.
- Conjugate directions: $d^n$ cycles through an $A$-orthogonal set of vectors.
Different descent methods: choice of functional $J$

- If $A$ is SPD the most common choice is

$$J(x) = \frac{1}{2} x^t A x - x^t b$$

- Minimum residual methods: for general $A$,

$$J(x) = \frac{1}{2} \langle b - Ax, b - Ax \rangle.$$

- Various combinations such as residuals plus updates:

$$J(x) = \frac{1}{2} \langle b - Ax, b - Ax \rangle + \frac{1}{2} \langle x - x^n, x - x^n \rangle.$$
Text, 265, 268
Topics

The connection to optimization

Descent Methods

Application to Stationary Iterative Methods

Parameter selection

The Steepest Descent Method
Stationary Iterative Methods

- \( A = M - N \)
- Iterative method \( Mx^{n+1} = b + Nx^n \)
- Equivalently \( M(x^{n+1} - x^n) = b - Ax^n \)
Examples

- FOR: $M = \rho I, \; N = \rho I - A$
- Jacobi: $M = \text{diag}(A)$
- Gauss-Seidel: $M = D + L$ (lower triangular part of $A$).
Householder lemma (269)

Let $A$ be SPD
Let $x^n$ be given by $M(x^{n+1} - x^n) = b - Ax^n$
Let $e^n = x - x^n$ then

$$\langle e^n, Ae^n \rangle - \langle e^{n+1}, Ae^{n+1} \rangle = \langle (x^{n+1} - x^n), P(x^{n+1} - x^n) \rangle$$

where $P = M + M^T - A$. 
Convergence of FOR, Jacobi, GS and SOR

Corollary 270
For $A$ SPD,
$P$ SPD $\implies$ convergence is monotonic in the $A$ norm:

$$\|e_n\|_A > \|e_{n+1}\|_A > \ldots \rightarrow 0 \quad \text{as} \quad n \rightarrow 0.$$
Convergence of FOR, Jacobi, GS and SOR

Corollary 270
For $A$ SPD,
$P$ SPD $\implies$ convergence is monotonic in the $A$ norm:

$$\|e_n\|_A > \|e_{n+1}\|_A > \ldots \rightarrow 0 \text{ as } n \rightarrow 0.$$ 

Householder lemma in this case gives

$$\langle e^n, Ae^n \rangle - \langle e^{n+1}, Ae^{n+1} \rangle = \langle (x^{n+1} - x^n), P(x^{n+1} - x^n) \rangle,$$
so

$$\langle e^n, Ae^n \rangle > \langle e^{n+1}, Ae^{n+1} \rangle,$$ or

$$\|e^{n+1}\|_A < \|e^n\|_A.$$
Proof

- $\|e^n\|_A \geq 0$ and monotone decreasing, so has a limit
- Cauchy criterion:
  
  $$(\langle e^n, Ae^n \rangle - \langle e^{n+1}, Ae^{n+1} \rangle) \to 0.$$

- Householder gives
  
  $\|x_{n+1} - x_n\|_P \to 0.$

- So that
  
  $M (x^{n+1} - x^n) = b - Ax^n \to 0$
Monotone convergence

1. FOR converges monotonically in $\| \cdot \|_A$ if

   $$P = M + M^t - A = 2\rho I - A > 0 \quad \text{if} \quad \rho > \frac{1}{2} \lambda_{\text{max}}(A).$$

2. Jacobi converges monotonically in the $A$-norm if $\text{diag}(A) > \frac{1}{2} A$.

3. Gauss-Seidel converges monotonically in the $A$ norm for SPD $A$ in all cases.

4. SOR converges monotonically in the $A$-norm if $0 < \omega < 2$. 

Proofs

1. For Jacobi, \( M = \text{diag}(A) \), so

\[
P = M + M^t - A = 2\text{diag}(A) - A > 0 \text{ if } \text{diag}(A) > \frac{1}{2}A.
\]
Proofs

1. For Jacobi, $M = \text{diag}(A)$, so

$$P = M + M^t - A = 2\text{diag}(A) - A > 0 \text{ if } \text{diag}(A) > \frac{1}{2}A.$$ 

2. For GS, $A = D + L + L^T$ (since $A$ SPD)

$$P = M + M^T - A = (D + L) + (D + L)^T - A =$$

$$D + L + D + L^T - (D + L + L^T) = D > 0$$
Proofs

1. For Jacobi, \( M = \text{diag}(A) \), so

\[
P = M + M^t - A = 2\text{diag}(A) - A > 0 \text{ if } \text{diag}(A) > \frac{1}{2}A.
\]

2. For GS, \( A = D + L + L^T \) (since \( A \) SPD)

\[
P = M + M^T - A = (D + L) + (D + L)^T - A = D + L + D + L^T - (D + L + L^T) = D > 0
\]

3. For SOR,

\[
P = M + M^t - A = M + M^t - (M - N) = M^t + N
\]

\[
= \omega^{-1} D + L^t + \frac{1 - \omega}{\omega} D - U = \frac{2 - \omega}{\omega} D > 0,
\]

for \( 0 < \omega < 2 \)
Proofs

1. For Jacobi, $M = \text{diag}(A)$, so

$$P = M + M^t - A = 2\text{diag}(A) - A > 0 \text{ if } \text{diag}(A) > \frac{1}{2} A.$$  

2. For GS, $A = D + L + L^T$ (since $A$ SPD)

$$P = M + M^T - A = (D + L) + (D + L)^T - A =$$

$$D + L + D + L^T - (D + L + L^T) = D > 0$$

3. For SOR,

$$P = M + M^t - A = M + M^t - (M - N) = M^t + N$$

$$= \omega^{-1} D + L^t + \frac{1 - \omega}{\omega} D - U = \frac{2 - \omega}{\omega} D > 0,$$

for $0 < \omega < 2$

4. Convergence of FOR in the $A$-norm is left as an exercise.
Text, Exercises 275, 278
In Exercise 278, you are to:

1. Give details of the proof that $P > 0$ implies convergence of FOR.
2. Prove the Householder lemma.
   Hint: expand the right side using $P = M + M^T - A$ and the identities $x^{n+1} - x^n = e^n - e^{n+1}$ and $\langle y, Wz \rangle = \langle W^T y, z \rangle$. 
Topics

The connection to optimization

Descent Methods

Application to Stationary Iterative Methods

Parameter selection

The Steepest Descent Method
Picking $\rho$ for FOR

- FOR: $\rho_{\text{optimal}} = (\lambda_{\text{max}} + \lambda_{\text{min}})/2$
- This gives fastest overall convergence for fixed $\rho$
- Hard to calculate.
- What about $\rho^n$ = “Best possible for this step”
Variable $\rho$

Given $x^0$ and a maximum number of iterations, $\text{itmax}$:

\begin{verbatim}
for n=0:itmax
    Compute $r^n = b - Ax^n$
    Compute $\rho^n$ via a few auxiliary calculations
    $x^{n+1} = x^n + (\rho^n)^{-1} r^n$
    if converged, exit, end
end
\end{verbatim}
Residuals $r^n = b - A \times x^n$ for FOR satisfy

$$r^{n+1} = r^n - \rho^{-1} Ar^n.$$

Proof:

$$x^{n+1} = x^n - \rho^{-1} r^n.$$

Multiply by “$-A$” and add $b$ to both sides

$$b - A x^{n+1} = b - A x^n - \rho^{-1} Ar^n,$$
Text, Exercise 281
Easy! Preparation for following work.
Option 1 Residual Minimization

Pick $\rho^n$ to minimize $\|r^{n+1}\|^2$:

$$\|r^{n+1}\|^2 = \|r^n - \rho^{-1} Ar^n\|^2 = \langle r^n - \rho^{-1} Ar^n, r^n - \rho^{-1} Ar^n \rangle.$$ 

Since $r^n$ is fixed, this is a simple function of $\rho$

$$\tilde{J}(\rho) = \langle r^n - \rho^{-1} Ar^n, r^n - \rho^{-1} Ar^n \rangle$$

or

$$\tilde{J}(\rho) = \langle r^n, r^n \rangle - 2 \langle r^n, Ar^n \rangle \rho^{-1} - \langle Ar^n, Ar^n \rangle \rho^{-2}.$$

Taking $\tilde{J}'(\rho) = 0$ and solving for $\rho = \rho_{\text{optimal}}$ gives

$$\rho_{\text{optimal}} = \frac{\langle Ar^n, Ar^n \rangle}{\langle r^n, Ar^n \rangle}.$$ 

The cost of using this optimal value at each step: two extra dot products per step.
Minimize $J(x^{n+1})$:

$$\phi(\rho) = J(x^n + \rho^{-1}r^n) = \frac{1}{2} \langle x^n + \rho^{-1}r^n, A(x^n + \rho^{-1}r^n) \rangle - \langle x^n + \rho^{-1}r^n, b \rangle$$

Expanding, setting $\phi'(\rho) = 0$ and solving, as before, gives

$$\rho_{\text{optimal}} = \frac{\langle r^n, Ar^n \rangle}{\langle r^n, r^n \rangle}.$$ 

- Option 2 requires SPD $A$
- Faster than Option 1
Algorithm for Option 2

Given $x^1$
the matrix $A$
a maximum number of iterations, $\text{itmax}$:

$$r^1 = b - Ax^1$$

for $n=1:\text{itmax}$

$$\rho^n = \langle Ar^n, r^n \rangle / \langle r^n, r^n \rangle$$

$$x^{n+1} = x^n + (1/\rho^n) r^n$$

if satisfied, exit, end

$$r^{n+1} = b - Ax^{n+1}$$

end
Homework

Text, Exercise 283. Implement and test Options 1 and 2.
The “normal” equations

- $Ae = r$ so $\|r\|_2^2 = \|Ae\|_2^2 = \|e\|_{A^TA}^2$
- $A^TA$ is SPD, so minimizing $\|e\|_{A^TA}^2$ is minimizing

$$(A^TA)x = A^Tb.$$  

- Both bandwidth and condition number of $A^TA$ are much larger than $A$, but $A^TA$ is SPD.
Let $A$ be $N \times N$ and invertible. Then $A^T A$ is SPD.
If $A$ is SPD then $\lambda(A^T A) = \lambda(A)^2$ and

$$\text{cond}_2(A^T A) = [\text{cond}_2(A)]^2.$$
Condition number of normal equations (284)

Let $A$ be $N \times N$ and invertible.
Then $A^T A$ is SPD.
If $A$ is SPD then $\lambda(A^T A) = \lambda(A)^2$ and

$$
cond_2(A^T A) = [cond_2(A)]^2.
$$

Proof

- Symmetry: $(A^T A)^T = A^T A^{TT} = A^T A$
- Positivity: $\langle x, AA^T x \rangle = \langle A^T x, A^T x \rangle = \|A^T x\|^2 > 0$
- Condition no.: For $A$ SPD

$$
cond_2(A^T A) = cond_2(A^2) = \frac{\lambda_{\max}(A^2)}{\lambda_{\min}(A^2)} = \left(\frac{\lambda_{\max}(A)}{\lambda_{\min}(A)}\right)^2 = (cond_2(A))^2.
$$
Condition number of normal equations (284)

Let \( A \) be \( N \times N \) and invertible. Then \( A^T A \) is SPD.
If \( A \) is SPD then \( \lambda(A^T A) = \lambda(A)^2 \) and

\[
\text{cond}_2(A^T A) = [\text{cond}_2(A)]^2.
\]

Proof

- **Symmetry:** \((A^T A)^T = A^T A^{TT} = A^T A\)
- **Positivity:** \(\langle x, AA^T x \rangle = \langle A^T x, A^T x \rangle = \|A^T x\|^2 > 0\)
- **Condition no.:** For \( A \) SPD

\[
\text{cond}_2(A^T A) = \text{cond}_2(A^2) = \frac{\lambda_{\text{max}}(A^2)}{\lambda_{\text{min}}(A^2)} = \left(\frac{\lambda_{\text{max}}(A)}{\lambda_{\text{min}}(A)}\right)^2 = (\text{cond}_2(A))^2.
\]

- Squaring the condition number greatly increases number of iterations
Topics

The connection to optimization

Descent Methods

Application to Stationary Iterative Methods

Parameter selection

The Steepest Descent Method
Steepest descent

- Reduce $J$ as much as possible in gradient direction
- Option 2
Steepest descent algorithm

Given
- SPD $A$
- $x^0, r^0 = b - Ax^0$
- maximum number of iterations $\text{itmax}$

for $n=0:$ $\text{itmax}$
  $r^n = b - Ax^n$
  $\alpha_n = \langle r^n, r^n \rangle / \langle r^n, Ar^n \rangle$
  $x^{n+1} = x^n + \alpha_n r^n$
  if converged, exit, end
end

$\alpha_n$ was called $1/\rho^n$ earlier.
Example: It is still slow!

- $N = 2$, $x = (x_1, x_2)^T$.
- $A = \begin{bmatrix} 2 & 0 \\ 0 & 50 \end{bmatrix}$, $b = \begin{bmatrix} 2 \\ 0 \end{bmatrix}$

$$J(\vec{x}) = \frac{1}{2} [x_1, x_2] \begin{bmatrix} 2 & 0 \\ 0 & 50 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} - [x_1, x_2] \begin{bmatrix} 2 \\ 0 \end{bmatrix}$$

$$= \frac{1}{2} (2x_1^2 + 50x_2^2) - 2x_1 = x_1^2 - 2x_1 + 25x_2^2 + 1 - 1$$

$$= \frac{(x_1 - 1)^2}{1^2} + \frac{x_2^2}{(\frac{1}{5})^2} - 1.$$
Steepest descent illustration

\[ x^0 = (11, 1) \]
\[ x^4 = (6.37, 0.54) \]
\[ \text{limit} = (1, 0) \]
Let $A$ be SPD and $\kappa = \lambda_{\text{max}}(A)/\lambda_{\text{min}}(A)$. The steepest descent method converges to the solution of $Ax = b$ for any $x^0$. The error $x - x^n$ satisfies

$$\|x - x^n\|_A \leq \left(\frac{\kappa - 1}{\kappa + 1}\right)^n \|x - x^0\|_A$$

and

$$J(x^n) - J(x) \leq \left(\frac{\kappa - 1}{\kappa + 1}\right)^n (J(x^0) - J(x)).$$
SD is no faster than FOR (287)

Let $A$ be SPD and $\kappa = \lambda_{\text{max}}(A)/\lambda_{\text{min}}(A)$. The steepest descent method converges to the solution of $Ax = b$ for any $x^0$. The error $x - x^n$ satisfies

$$
\|x - x^n\|_A \leq \left( \frac{\kappa - 1}{\kappa + 1} \right)^n \|x - x^0\|_A
$$

and

$$
J(x^n) - J(x) \leq \left( \frac{\kappa - 1}{\kappa + 1} \right)^n (J(x^0) - J(x)).
$$

But no eigenvalue estimates are required!
Proof

- $\langle x - x^n, A(x - x^n) \rangle = \langle x, b \rangle + 2J(x^n)$
- Minimizing $J$ is the same as minimizing $\|x - x^n\|_A$
- Steepest descent reduces $J$ maximally
- FOR reduces $\|x - x^n_{FOR}\|_A$ with known bound
- Doing one iteration of each, starting from same $x^{n-1}$,

$$\|x - x^n\|_A \leq \|x - x^n_{FOR}\|_A \leq \left(\frac{\kappa - 1}{\kappa + 1}\right) \|x - x^{n-1}\|_A$$
Theorem 288: inequality is sharp!

- $Ax = b$
- If $x^0 = x - (\lambda_2 \phi_1 \pm \lambda_1 \phi_2)$ where $\phi_1$ and $\phi_2$ are the eigenvectors of $\lambda_1 = \lambda_{\min}(A)$ and $\lambda_2 = \lambda_{\max}(A)$ respectively
- Then the convergence rate is precisely

$$\frac{\lambda_2 - \lambda_1}{\lambda_2 + \lambda_1} = \frac{\kappa - 1}{\kappa + 1}$$
Proof

\[ x^0 = x - (\lambda_2 \phi_1 \pm \lambda_1 \phi_2) \]
\[ e^0 = \lambda_2 \phi_1 \pm \lambda_1 \phi_2 \]
Proof

\[ x^0 = x - (\lambda_2 \phi_1 \pm \lambda_1 \phi_2) \]
\[ e^0 = \lambda_2 \phi_1 \pm \lambda_1 \phi_2 \]
\[ r^0 = Ae^0 = -\lambda_1 \lambda_2 \phi_1 \mp \lambda_1 \lambda_2 \phi_2 \]
Proof

\[ x^0 = x - (\lambda_2 \phi_1 \pm \lambda_1 \phi_2) \]
\[ e^0 = \lambda_2 \phi_1 \pm \lambda_1 \phi_2 \]
\[ r^0 = Ae^0 = -\lambda_1 \lambda_2 \phi_1 \mp \lambda_1 \lambda_2 \phi_2 \]
\[ Ar^0 = -\lambda_1^2 \lambda_2 \phi_1 \mp \lambda_1 \lambda_2^2 \phi_2 \]
Proof

\[ x^0 = x - (\lambda_2 \phi_1 \pm \lambda_1 \phi_2) \]
\[ e^0 = \lambda_2 \phi_1 \pm \lambda_1 \phi_2 \]
\[ r^0 = A e^0 = -\lambda_1 \lambda_2 \phi_1 \mp \lambda_1 \lambda_2 \phi_2 \]
\[ Ar^0 = -\lambda_1^2 \lambda_2 \phi_1 \mp \lambda_1 \lambda_2^2 \phi_2 \]
\[ \alpha_0 = \langle r^0, r^0 \rangle / \langle r^0, Ar^0 \rangle \]
\[ = \frac{\lambda_1^2 \lambda_2^2 + \lambda_1^2 \lambda_2^2}{\lambda_1^3 \lambda_2^2 + \lambda_1^2 \lambda_2^3} = \frac{2}{\lambda_1 + \lambda_2} \]
Proof

\[ x^0 = x - (\lambda_2 \phi_1 \pm \lambda_1 \phi_2) \]

\[ e^0 = \lambda_2 \phi_1 \pm \lambda_1 \phi_2 \]

\[ r^0 = A e^0 = -\lambda_1 \lambda_2 \phi_1 \mp \lambda_1 \lambda_2 \phi_2 \]

\[ Ar^0 = -\lambda_1^2 \lambda_2 \phi_1 \mp \lambda_1 \lambda_2^2 \phi_2 \]

\[ \alpha_0 = \langle r^0, r^0 \rangle / \langle r^0, Ar^0 \rangle \]

\[ = \frac{\lambda_1^2 \lambda_2^2 + \lambda_1^2 \lambda_2^2}{\lambda_1^3 \lambda_2^2 + \lambda_1^2 \lambda_2^3} = \frac{2}{\lambda_1 + \lambda_2} \]

\[ x^1 = x^0 + \alpha r^0 \]

\[ = x - \left( \lambda_2 - \frac{2}{\lambda_1 + \lambda_2} \lambda_1 \lambda_2 \right) \phi_1 - \left( \mp \lambda_1 \mp \frac{2}{\lambda_1 + \lambda_2} \lambda_1 \lambda_2 \right) \phi_2 \]

\[ = x - \lambda_2 \left( \frac{\lambda_2 - \lambda_1}{\lambda_2 + \lambda_1} \right) \phi_1 - \lambda_1 \left( \mp \lambda_1 \mp \frac{2}{\lambda_1 + \lambda_2} \lambda_1 \lambda_2 \right) \phi_2 \]

\[ = x - \frac{\lambda_2 - \lambda_1}{\lambda_2 + \lambda_1} (\lambda_2 \phi_1 \mp \lambda_1 \phi_2) \]
Text, Exercise 291
Exercise 290: Convection-Diffusion Equation

\[-\epsilon \Delta u + c \cdot \nabla u = f\] inside \( \Omega = [0, 1] \times [0, 1] \)

\[c = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \]

\[u = g \text{ on } \partial \Omega\]
Exercise 290: Convection-Diffusion Equation

\[-\epsilon \Delta u + c \cdot \nabla u = f \text{ inside } \Omega = [0, 1] \times [0, 1]\]

\[c = \begin{bmatrix} 1 \\ 0 \end{bmatrix}\]

\[u = g \text{ on } \partial \Omega\]

- The convection term is discretized as

\[c \cdot \nabla u = \frac{\partial u}{\partial x} \approx \frac{u(i+1,j) - u(i-1,j)}{2h}\]
Comments about CDE

- $c$ is like a wind, blowing the solution to the right.
- Interesting behavior for $\epsilon < \|c\|$
Comments about CDE

- $c$ is like a wind, blowing the solution to the right.
- Interesting behavior for $\epsilon < \|c\|$
- $Pe= ”cell Péclet number”$
- *Discretization* becomes unstable as $Pe$ becomes large.
Modify `jacobi2d.m`

- Add in convection term.
- Ends up with a factor of \( h \) in the numerator.
- Debug with \( f = 1, \ g = x - y \).
  - Exact solution is \( u = g \).
  - Also exact discrete solution!
  - Diffusion is debugged already.
- Debug with \( N = 5 \)
  - Can see the solution on the screen.
  - \( h = 0.2 \) so can recognize solution values on sight.
- Start from exact: must get answer in 1 iteration.
- Start from exact, run many iterations: must stay exact.
- Start from zero, get exact?
  - If not, program in exact solution, print norm of error at each iteration.
Modify `jacobi2d.m`

- Add in convection term.
- Ends up with a factor of $h$ in the numerator
- Debug with $f = 1$, $g = x - y$.
  - Exact solution is $u = g$
  - Also exact discrete solution!
  - Diffusion is debugged already
- Debug with $N = 5$
  - Can see the solution on the screen
  - $h = 0.2$ so can recognize solution values on sight
- Start from exact: must get answer in 1 iteration.
- Start from exact, run many iterations: must stay exact
- Start from zero, get exact?
  - If not, program in exact solution, print norm of error at each iteration.
- Double-check original code!
“Real” problem

- $N = 50$, $f = x + y$, $g = 0$
- Jacobi, steepest descent, residual minimization
- $\epsilon = 1.0, 0.1, 0.01, 0.001$
Solution, $\epsilon = 1.0$
Solution, $\epsilon = 0.1$
Solution, $\epsilon = 0.01$
Solution, $\epsilon = 0.001$
Iterative methods

<table>
<thead>
<tr>
<th>$\epsilon$</th>
<th>Jacobi</th>
<th>Steepest descent</th>
<th>Residual minimization</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.0</td>
<td>3318</td>
<td>3246</td>
<td>3281</td>
</tr>
<tr>
<td>0.1</td>
<td>17161</td>
<td>1652</td>
<td>1690</td>
</tr>
<tr>
<td>0.01</td>
<td>13235</td>
<td>130</td>
<td>113</td>
</tr>
<tr>
<td>0.001</td>
<td>15735</td>
<td>diverges</td>
<td>1663</td>
</tr>
</tbody>
</table>