Iterative Solvers

Last Time: Eigenvalue calculations

1. Power method:
   \[ A^k v \approx \lambda_i^k v_i, \quad |\lambda_i| > |\lambda_2| \ldots \]
   converges at a rate \( O\left(\frac{1}{\lambda_i^k}\right) \).

2. To increase convergence rate: Shift
   \( A - \sigma I \) has eigenvalues equal to \( \lambda_i - \sigma \), \( \lambda_2 - \sigma \ldots \)
   pick \( \sigma \) to increase convergence rate.

Ex:
\[ \lambda_j = 1, 2, \ldots, 8, 9, 10 \]
\[ \lambda_1 = 10, \quad \lambda_2 = 9 \]

Power method converges at a rate of \( \left(\frac{9}{10}\right)^k \).

Shift by \( \sigma = 1 \), \( \lambda_1 = 9, \quad \lambda_2 = 8 \ldots \lambda_{10} = 0 \)

Rate: \( \left(\frac{9}{10}\right)^k \cdot \left(\frac{9}{10}\right)^k \).

Best shift: Shift such that \( |\lambda_2| = |\lambda_{10}| \)

\[ \sigma = \frac{\lambda_1}{2} = 5 \]
\[ \lambda_1 = 5 \]
\[ \lambda_2 = 4 \]
\[ \lambda_{10} = -4 \]

\( \text{(Eigenvalues did not change)} \)
To find interior eigenvalues, the inverse iteration with shift is required:

\[(A - sI)^{-1}\] has eigenvalues \(\frac{1}{\lambda_1 - s}, \ldots, \frac{1}{\lambda_{10} - s}\).

Example: for \(s = 3.2\), \(\tilde{\lambda}_j\) are

\[\tilde{\lambda}_1 = \frac{1}{10 - 3.2}, \quad \tilde{\lambda}_2 = \frac{1}{9 - 3.2}, \quad \tilde{\lambda}_{10} = \frac{1}{1 - 3.2}\]

The largest absolute value \(\tilde{\lambda}\) is \(\tilde{\lambda}_8 = \frac{1}{3 - 3.2} = \frac{1}{-0.2} = -5\).

\[\tilde{\lambda}_7 = \frac{1}{4} = 1.25, \quad \tilde{\lambda}_9 = \frac{1}{1.2} = -\frac{5}{6}\]

\(\Rightarrow\) \(\tilde{\lambda}_j\)'s can be estimated with power method + deflate...

How do we compute \(\left((A - sI)^{-1}\right)^k \vec{w}_0\)?

\(\Rightarrow\) Solve \((A - sI) \vec{w}_k + \vec{w}_0 = \vec{w}_0\)

Into: Iterative solver.

Cost of Gaussian elimination: \(O(n^3)\)

" " matrix-vector-mult: \(O(n^2)\) \(\Rightarrow\) maybe we can iterate \(k\) times

\(\Rightarrow\) \(O(kn^2)\).
Another reason we want iterative solvers: storage.

\[
A = \begin{pmatrix}
2 & -1 \\
-1 & 2 \\
-1 & 2 & -1 \\
-1 & 2
\end{pmatrix}
\]

Storage cost for \( A \): \( O(3n^2) \)

\( A^{-1} \) is dense! (check in Matlab for yourself)

storage: \( O(n^2) \).

If \( A \) is \( n \times n \), and contains \( O(n) \) or \( O(n \log n) \) entries, we say that it is **sparse**.

**Simple Iteration**

Suppose that you have a matrix \( M \) such that \( M^{-1}A = I \) and \( M\vec{x} = \vec{y} \) is easy to solve.

Then: To solve \( A\vec{x} = \vec{b} \):

1. Start with guess \( \vec{x}_0 \).
2. Compute residual \( \vec{r}_0 = \vec{b} - A\vec{x}_0 \).
3. Solve \( M\vec{z}_0 = \vec{r}_0 \), then \( \vec{x}_0 = M^{-1}(\vec{b} - A\vec{x}_0) \).
4. Set \( \vec{x}_1 = \vec{x}_0 + \frac{1}{\alpha} \vec{z}_0 \).

\( \alpha = \frac{\vec{r}_0^T \vec{r}_0}{\vec{z}_0^T \vec{z}_0} \).
And repeat:

$k = 1, 2, \ldots$

Set $\tilde{x}_k = \frac{1}{2} (\tilde{x}_{k-1} + \tilde{z}_{k-1})$

Compute $\tilde{r}_k = \frac{1}{2} (\tilde{b} - \tilde{A}\tilde{x}_k)$

If $||\tilde{r}_k||$ is small, stop

Else solve $M\tilde{z}_k = \tilde{r}_k$

$M$ is referred to as a left preconditioner.

How do we choose $M$?

Ex: $M = \text{Lower triangle of } A$. (Gauss-Seidel iterations)

Easy to solve, $O(n^2)$ flops by forward substitution

Convergence of Iterative Methods

Consider the matrix splitting: $A = M - N$.

$A\tilde{x} = \tilde{b}$

$(M-N)\tilde{x} = \tilde{b}$

$\tilde{x} = M^{-1}N\tilde{x} + M^{-1}\tilde{b}$ is similar to fixed point iteration with $q(x) = M^{-1}N x + b M^{-1}$.
This is, in fact, equivalent to the earlier iteration method:

\[ \hat{x}_k = \hat{x}_{k-1} + \frac{1}{2} \hat{z}_{k-1} \]

\[ \left( M^* \frac{1}{2} \hat{x}_k = \hat{x}_k \right) \]

\[ = \left[ 1 \right] \hat{x}_{k-1} + M^* \left[ b - A \hat{x}_{k-1} \right] \]

\[ = \left[ I - M^* A \right] \hat{x}_{k-1} + M^* b, \]

and thus:

\[ A = M - N \]

\[ M^{-1} N = I - M^{-1} A = M^{-1} (M - A) \]

Same as before.

So when does \( \hat{x}_k = \left( I - M^* A \right) \hat{x}_{k-1} \) converge?

At step \( k \), the error is:

\[ \hat{e}_k = A^{-1} b - \hat{x}_k \]

\[ \hat{x}_k = \hat{x}_{k-1} + \hat{z}_{k-1} \]

\[ - \hat{\chi}_k + A^{-1} b = - \hat{x}_{k-1} + A^{-1} b - \hat{z}_{k-1} \]

\[ \hat{e}_k = \hat{e}_{k-1} + M^{-1} \hat{v}_{k-1} \]

\[ \hat{e}_k = \hat{e}_{k-1} + M^{-1} \hat{v}_{k-1} \]

\[ \hat{e}_k = A \left( A^{-1} b - \hat{x}_k \right) \]

\[ = A \hat{e}_k \]

\[ \hat{e}_k = \hat{e}_{k-1} + M^{-1} A \hat{e}_{k-1} \]

\[ = \left( I - M^* A \right) \hat{e}_{k-1} \]
This implies that
\[ \hat{e}_k = (I - MA)^k \hat{e}_0. \]
\[ \Rightarrow \| \hat{e}_k \| \leq \| (I - MA)^k \| \| \hat{e}_0 \| \]
for any vector norm and the induced matrix norm.

**Theorem:** \( \| \hat{e}_k \| \to 0 \) and \( \hat{x}_k \to \bar{x} \) if and only if
\[ \lim_{k \to 0} \| (I - MA)^k \| = 0 \]
for every initial \( \hat{x}_0 \).

(Unsurprising)

**Example** Recall \( \| A \|_2 \).
\[ \| A \|_2 = \| A^T A \|^{1/2} \]
\[ \Rightarrow \| A \|_2 = \max \{ |\lambda_j| \} \]

Therefore we need the eigenvalues of \( I - MA \) to be \( \leq 1 \) in absolute value.
Definition: Spectral radius:

For an \( n \times n \) matrix \( G \),

\[
\rho(G) = \max_j |\lambda_j|
\]

Theorem: Let \( G \) be \( n \times n \) matrix. Then \( \lim_{k \to \infty} G^k = 0 \)

if and only if \( \rho(G) \leq 1 \).

Idea: If \( G = V \Lambda V^{-1} \), then \( G^k = V \Lambda^k V^{-1} \)

\( \to 0 \) iff elements \( \leq 1 \).

\( \Rightarrow \) for convergence, we need \( \rho(I - M^2A) \leq 1 \).

(And if \( \| (I - M^2A)^k \| \to 0 \), then \( \rho(I - M^2A) \leq 1 \)).

Special case: symmetric positive definite matrix.

No preconditioner, consider only multiple of \( A \), and the

graphical iteration.

\[
\begin{align*}
X_0 & = X_0 - 2A^2x_0 \\
X_1 & = X_0 + C_0 \frac{1}{2}
\end{align*}
\]

\( \vdots \)

\[
\begin{align*}
X_2 & = X_1 + C_1 \frac{1}{2} \\
\vdots
\end{align*}
\]
Conjugate Gradient for RSPD matrix

Def:
\( \tilde{\mathbf{w}}, \tilde{\mathbf{v}} \) are conjugate w.r.t. to \( \mathbf{A} \) if
\[ (\tilde{\mathbf{w}}, \mathbf{A}\tilde{\mathbf{v}}) = 0 \] \( \text{(the A-inner product)} \)

Note: \( (\tilde{\mathbf{w}}, \mathbf{A}\tilde{\mathbf{v}}) \) is only an inner product if \( \mathbf{A} \) is symmetric and positive definite.

Let \( \{\tilde{\mathbf{v}}_1, ..., \tilde{\mathbf{v}}_n\} \) be a set of \( n \) mutually \( \mathbf{A} \)-conjugate vectors. Then they form a basis for \( \mathbb{R}^n \) (if \( \mathbf{A} \) is invertible).

Any \( \tilde{\mathbf{x}} \) can be written as
\[ \tilde{\mathbf{x}} = \sum \alpha_j \tilde{\mathbf{v}}_j \]

Therefore:
\[ \mathbf{A}\tilde{\mathbf{x}} = \sum \alpha_j \mathbf{A}\tilde{\mathbf{v}}_j = \sum \alpha_j (\tilde{\mathbf{v}}_j, \mathbf{A}\tilde{\mathbf{v}}_j) \tilde{\mathbf{v}}_j \]
\[ = \alpha_1 (\tilde{\mathbf{v}}_1, \mathbf{A}\tilde{\mathbf{v}}_1) = (\tilde{\mathbf{v}}_1, b) \]
\[ = \mathbf{A}\tilde{\mathbf{x}} = (\tilde{\mathbf{v}}_1, b) \]
\[ \Rightarrow \tilde{\mathbf{x}} = \frac{(\tilde{\mathbf{v}}_1, b)}{\|\tilde{\mathbf{v}}_1\|_A^2} \quad \text{Direct solution} \]
To do this iteratively, is there a way to select $\hat{\mathbf{p}}_1, \ldots, \hat{\mathbf{p}}_k$ so that only a few inner products need to be calculated? (as an approximation to $\mathbf{x}$ s.t. $\mathbf{x}^* \mathbf{b} = \mathbf{A}^T \mathbf{x}$ is small).

Observation: The solution to $\mathbf{A} \hat{\mathbf{x}} = \hat{\mathbf{b}}$ is also the unique minimizer of the function:

$$ f(\hat{\mathbf{x}}) = \frac{1}{2} (\hat{\mathbf{x}}, \mathbf{A} \hat{\mathbf{x}}) - (\hat{\mathbf{b}}, \hat{\mathbf{x}}) $$

(i.e. $\frac{\partial f}{\partial \mathbf{x}_j} = 0$ when $\mathbf{A} \hat{\mathbf{x}} = \hat{\mathbf{b}}$).

Newton would suggest moving in the direction of $\nabla f(\mathbf{x}_0)$.

$$ \nabla f(\mathbf{x}) = \mathbf{A} \hat{\mathbf{x}} - \hat{\mathbf{b}}. $$

Let $\mathbf{p}_0 = \mathbf{b} - \mathbf{A} \hat{\mathbf{x}} = \mathbf{v}_0 = -\nabla f(\mathbf{x}_0).$

We want the following values $\mathbf{p}_1, \mathbf{p}_2, \ldots$ to be $\mathbf{A}$-conjugate:

Let $\mathbf{p}_k = \mathbf{b} - \mathbf{A} \hat{\mathbf{x}}_k = -\nabla f(\mathbf{x}_k)$

$$ \Rightarrow \mathbf{p}_k = \hat{\mathbf{p}}_k - \sum_{l \neq k} \frac{(\mathbf{p}_l, \mathbf{A} \hat{\mathbf{x}}_l)}{(\mathbf{p}_l, \mathbf{p}_l)} \mathbf{p}_l $$

"$\mathbf{A}$-minor product"

"Gram-Schmidt"
In the direction of \( \tilde{\beta}_k \), the next iterate is:

\[
\tilde{x}_{k+1} = \tilde{x}_k + \alpha_k \tilde{\beta}_k
\]

with \( \alpha_k = \frac{(\tilde{\beta}_k, b)}{(\tilde{\beta}_k, A\tilde{\beta}_k)} = \frac{(\tilde{\beta}_k, \tilde{v}_{k-1} + A\tilde{x}_{k-1})}{(\tilde{\beta}_k, A\tilde{\beta}_k)} = \frac{(\tilde{\beta}_k, \tilde{v}_{k-1})}{(\tilde{\beta}_k, A\tilde{\beta}_k)}
\)

since \( (\tilde{\beta}_k, A\tilde{x}_{k-1}) \) are A-conjugate.

**Convergence Rate**

Then let \( \tilde{e}_k \) be the error on step \( k \) of CG. Then

\[
\frac{\|\tilde{e}_k\|_A}{\|\tilde{e}_0\|_A} \leq 2 \left( \frac{\nu R - 1}{\nu R + 1} \right)^k\]

\( K = \frac{\lambda_{\text{max}}}{\lambda_{\text{min}}} \) (condition number)

(for RSPP matrices)

Conjugate Gradient is a **Krylov Method**

It minimizes some norm of \( \|\tilde{x}_k\| \) in fact the

using \( \tilde{v}_0, A\tilde{v}_0, A^2\tilde{v}_0, \ldots \)

For nonsymmetric, arbitrary systems \( A\tilde{x} = \tilde{b} \), the

**GMRES** algorithm minimizes the 2 norm \( \|\tilde{x}_k\| \).
Idea: solve $Ax = b$ by choosing an approximate iterate $\tilde{x}_k$ that is of the form
$\tilde{x}_0 + \sum_{j=1}^{k} c_j A^j \tilde{v}_j$

minimize $\| \tilde{v}_k - \tilde{v}_0 - \sum_{j=1}^{k} c_j A^j \tilde{v}_j \|_2$ in the 2-norm.

(Use Gram-Schmidt) and an orthogonal basis for $\{ \tilde{v}_0, A\tilde{v}_0, A^2\tilde{v}_0, \ldots \}$