The Linear Algebra of Image Filtering

Katrina Palmer

www.mathcs.emory.edu/~kpalmer
Outline

1. Solving $Ax = b$ and matrix factorizations

2. Singular value decomposition (SVD)

3. Regularization
   (a) Filtering
   (b) Iterative
Solving $Ax = b$

Simple?

- $x = A^{-1}b$

- In Matlab: `>> x = A \ b`
Solving $Ax = b$

Simple?

- $x = A^{-1}b$

- In Matlab: `>> x = A \ b`

In general, it's not so simple!
$A$ and $b$ are known

Right hand side:
$A$ and $b$ are known

Right hand side:

$A^{-1}b$
Solving $Ax = b$

- "Iterative" Methods
  - $x_0 = \text{initial guess}$
  - compute approximations $x_1, x_2, \ldots$
  - $x_k \rightarrow x$

- "Direct" Methods
  Based on matrix factorizations
Some Matrix Factorizations

Gaussian Elimination

\[ A = LU \] (no pivoting)
\[ PA = LU \] (pivoting)
\[ A = LL^T \] (for SPD matrix)
Some Matrix Factorizations

Gaussian Elimination

\[ A = LU \] (no pivoting)
\[ PA = LU \] (pivoting)
\[ A = LL^T \] (for SPD matrix)

Others

\[ A = QR \] (for least squares)
\[ A = XX^{-1} \] (eigenvalues)
Singular Value Decomposition (SVD)

\[ A = U \Sigma V^T \]

where

\( U \) is orthogonal \((U^T U = I)\)

\( V \) is orthogonal \((V^T V = I)\)

\( \Sigma = \text{diag}(\sigma_1, \ldots, \sigma_n) \)

\[ \sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_n \]
SVD Properties

\[ A = U \Sigma V^T \]

- diagonal entries \( \sigma_i = \) singular values
- columns vectors \( v_i = \) right singular vectors
- columns vectors \( u_i = \) left singular vectors

- \( \text{rank}(A) = r \iff \sigma_1 \geq \cdots \geq \sigma_r > \sigma_{r+1} = \cdots = \sigma_n = 0 \)
SVD Analysis of $Ax = b$

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

\[
= (U\Sigma V^T)^{-1}b
\]

\[
= V\Sigma^{-1}U^Tb
\]

\[
= \sum_{i=1}^{n} \frac{u_i^Tb}{\sigma_i}v_i
\]
SVD Analysis of \( Ax = b \)

\[
\hat{x} = A^{-1}(b + \varepsilon) \\
= (U\Sigma V^T)^{-1}(b + \varepsilon) \\
= V\Sigma^{-1}U^T(b + \varepsilon) \\
= \sum_{i=1}^{n} \frac{u_i^T(b + \varepsilon)}{\sigma_i} v_i
\]
SVD Analysis of $Ax = b$

\[
\hat{x} = A^{-1}(b + \varepsilon)
\]

\[
= V\Sigma^{-1}U^T(b + \varepsilon)
\]

\[
= \sum_{i=1}^{n} \frac{u_i^T(b + \varepsilon)}{\sigma_i}v_i
\]

\[
= \sum_{i=1}^{n} \frac{u_i^Tb}{\sigma_i}v_i + \sum_{i=1}^{n} \frac{u_i^T\varepsilon}{\sigma_i}v_i
\]

\[
= x + \text{error}
\]
SVD Analysis of $Ax = b$

\[
\hat{x} = A^{-1}(b + \varepsilon)
\]

\[
= x + \sum_{i=1}^{n} \frac{u_i^T \varepsilon}{\sigma_i} v_i
\]

Observations:

- If $\sigma_i \approx 0$, error can be large.

- $\frac{\sigma_1}{\sigma_n} = \text{“condition number”}$

large $\Rightarrow$ small change in data can mean very poor $\hat{x}$

small $\Rightarrow$ can get good $\hat{x}$ even with small change in data
Class of Examples

First kind integral equations:

\[ b(t) = \int_{\Omega} a(t, s)x(s) ds \]
\[ b = Ax + \varepsilon \quad \text{(discrete, linear algebra problem)} \]

Applications:

- Geomagnetic prospecting.
- Tomography.
- Image Restoration.
Class of Examples

First kind integral equations:

\[ b(t) = \int_{\Omega} a(t, s)x(s)\,ds \]

\[ b = Ax + \varepsilon \quad \text{(discrete, linear algebra problem)} \]

Applications:

- Geomagnetic prospecting.
- Tomography.
- **Image Restoration.**
  \( x = \text{true image} \)
  \( b = \text{blurred, noisy image} \)
  \( A = \text{blurring matrix (point spread function)} \)
Properties of Integral Equations

- singular values cluster at 0  (i.e., $A$ is ill-conditioned)
- No gap to separate large/small singular values
- small singular values $\Rightarrow$ oscillating singular vectors

Therefore, can have large solution error:

$$\sum_{i=1}^{n} \frac{u_i^T \varepsilon}{\sigma_i} v_i$$
Regularization by Filtering

- Naive solution \( x = A^{-1}b \) corrupted by noise

- Regularization \( \Rightarrow \) mathematical trick to stabilize solution

\[
x_{\text{reg}} = \sum_{i=1}^{n} \phi_i \frac{u_i^T b}{\sigma_i} v_i
\]

where the “filter factors” satisfy

\[
\phi_i \approx \begin{cases} 
1 & \text{if } \sigma_i \text{ is large} \\
0 & \text{if } \sigma_i \text{ is small}
\end{cases}
\]
Two Well Known Regularization Methods

Truncated SVD

\[ x_{tsvd} = \sum_{i=1}^{n} \phi_i \frac{u_i^T b}{\sigma_i} v_i \quad \Leftrightarrow \quad \phi_i = \begin{cases} 1 & \text{for } \sigma_i \geq \tau \\ 0 & \text{for } \sigma_i < \tau \end{cases} \]

Tikhonov

\[ x_{tik} = \sum_{i=1}^{n} \phi_i \frac{u_i^T b}{\sigma_i} v_i \quad \Leftrightarrow \quad \phi_i = \frac{\sigma_i^2}{\sigma_i^2 + \alpha^2} \]
Image Restoration Example

The image size is $256 \times 256$ which makes $A 256^2 \times 256^2$
Iterative Regularization

Big Idea: Stop the iterative method early (at the 'best' reconstruction)
Properties of Iterative Methods

• Early iterations reconstruct large singular components.

• Later iterations reconstruct noise.

Note: Iterative methods can be thought of as filtering methods
Two Well Known Iterative Methods

- Steepest Descent
  - step direction is the negative gradient
  - known for slow convergence

- Conjugate Gradient Method
  - step directions are orthogonal
  - theoretically converges in $n$ steps
The filter factors for both SD and CG are $\sigma_i^2 P_k(\sigma_i^2)$ where

- **Steepest Descent**
  \[
P_k(\lambda) = P_{k-1}(\lambda) + \alpha_k (1 - \lambda P_{k-1}(\lambda))
  \]
  \[
P_{-1}(\lambda) = 0.
  \]

- **Conjugate Gradient**
  \[
P_k(\lambda) = \left(1 - \alpha_k \lambda + \frac{\alpha_k \beta_{k-1}}{\alpha_{k-1}}\right) P_{k-1}(\lambda) - \frac{\alpha_k \beta_{k-1}}{\alpha_{k-1}} P_{k-2} + \alpha_k ,
  \]
  \[
P_{-1}(\lambda) = 0 \text{ and } P_0(\lambda) = \alpha_0.
  \]
Problems:

- SD converges too slowly
- A good stopping criteria is needed for CG

Remedy for slow convergence:
Problems:

- SD converges too slowly

- A good stopping criteria is needed for CG

Remedy for slow convergence: **Precondition!**
Preconditioning

- Speeds up convergence

- Multiply both sides of $Ax = b$ by $P^{-1}$ to make the problem "nicer"
Preconditioning Ill-Posed Problems

Modify TSVD idea (Hanke, Nagy, Plemmons, ’93; Hanke, Nagy, ’96)

Let \( P_\tau = U\Sigma_\tau V^T \)

where

- \( \Sigma_\tau = \text{diag}(\sigma_1, \ldots, \sigma_k, 1, \ldots, 1) \)
- \( \sigma_{k-1} \geq \tau > \sigma_k \)
Preconditioning Ill-Posed Problems

Notice that the preconditioned system is:

\[ P_\tau^{-1} A = (U \Sigma \tau V^T)^{-1}(U \Sigma V^T) \]
\[ = V \Sigma^{-1}_\tau \Sigma V^T \]
\[ = V \Delta V^T \]

where \( \Delta = \text{diag}(1, \ldots, 1, \sigma_{k+1}, \ldots, \sigma_n) \)

That is,

- Large (good) singular values clustered at 1.
- Small (bad) singular values not clustered.
Preconditioning Ill-Posed Problems

Remaining questions:

1. How to choose truncation, $\tau$?

2. We can’t compute SVD, so now what?
Preconditioning Ill-Posed Problems

Remaining questions:

1. How to choose truncation, τ?
   
   Use GCV, L-curve, Picard condition

2. We can’t compute SVD, so now what?
Preconditioning Ill-Posed Problems

Remaining questions:

1. How to choose truncation, $\tau$?

   Use GCV, L-curve, Picard condition

2. We can’t compute SVD, so now what?

   For image restoration, can find good spectral approximations. (i.e.- Circulant approximations)
A more real example
Relative Errors

![Relative Errors Chart]

- **CGLS**
- **RNSD**
- **PCGILS**
- **PRNSD**

36
Restorations at the 39th iteration
• Linear algebra is everywhere!

• The SVD is a useful tool for filtering blurred images.

• Iterative Methods as filtering

• On-going work includes:
  – Iterative Regularization (analyzing filter factors)
  – Further development of software package *RestoreTools*.

www.mathcs.emory.edu/~kpalmer