EECS 275 Matrix Computation

Ming-Hsuan Yang

Electrical Engineering and Computer Science
University of California at Merced
Merced, CA 95344
http://faculty.ucmerced.edu/mhyang

Lecture 5
Overview

- Matrix properties via singular value decomposition (SVD)
- Geometric interpretation of SVD
- Applications
Reading

- Chapter 5 of *Numerical Linear Algebra* by Llyod Trefethen and David Bau
- Chapter 3 of *Matrix Computations* by Gene Golub and Charles Van Loan
- Chapter 3 of *Mathematical Modeling of Continuous Systems* by Carlo Tomasi
- Chapter 5 of *Matrix Analysis and Applied Linear Algebra* by Carl Meyer
Full and reduced SVD

- Let $A \in \mathbb{R}^{m \times n}$
- Reduced SVD: $A = \hat{U}\hat{\Sigma}\hat{V}^\top$, $\hat{U} \in \mathbb{R}^{m \times n}$, $\hat{\Sigma} \in \mathbb{R}^{n \times n}$ and $\hat{V} \in \mathbb{R}^{n \times n}$
- Full SVD: $A = U\Sigma V^\top$, $U \in \mathbb{R}^{m \times m}$, $\Sigma \in \mathbb{R}^{n \times n}$ and $V \in \mathbb{R}^{n \times n}$

Reduced SVD ($m \geq n$)

$$A = \hat{U}\hat{\Sigma}\hat{V}^\top$$

Full SVD ($m \geq n$)

$$A = U\Sigma V^\top$$
Uniqueness

- First note that $\sigma_1$ and $v_1$ can be uniquely determined by $\|A\|_2$
- Suppose in addition to $v_1$, there is another linearly independent vector $w$ with $\|w\|_2 = 1$ and $\|Aw\|_2 = \sigma_1$
- Define a unit vector $v_2$, orthogonal to $v_1$ as a linear combination of $v_1$ and $w$

$$v_2 = \frac{w - (v_1^\top w)v_1}{\|w - (v_1^\top w)v_1\|_2}$$

- Since $\|A\|_2 = \sigma_1$, $\|Av_2\|_2 \leq \sigma_1$, but this must be an equality, for otherwise $w = cv_1 + sv_2$ for some constants $c$ and $s$ with $|c|^2 + |s|^2 = 1$, we would have $\|Aw\| < \sigma_1$
- $v_2$ is a second right singular vector of $A$ corresponding to $\sigma_1$
- Once $\sigma_1$, $v_1$, and $v_1$ are determined, the remainder of SVD is determined by the action of $A$ on the space orthogonal to $v_1$
- Since $v_1$ is unique up to a sign, the orthogonal space is unique defined and so are the remaining singular values
Matrix properties via SVD

Theorem

The rank of $A$ is $r$, the number of nonzero singular values.

Proof.

The rank of a diagonal matrix is equal to the number of its nonzero entries, and in SVD, $A = U\Sigma V^\top$ where $U$ and $V$ are of full rank. Thus, $\text{rank}(A) = \text{rank}(\Sigma) = r$

Theorem

$\|A\|_2 = \sigma_1$, and $\|A\|_F = \sqrt{\sigma_1^2 + \cdots + \sigma_r^2}$

Proof.

As $U$ and $V$ are orthogonal, $A = U\Sigma V^\top$, $\|A\|_2 = \|\Sigma\|_2$. By definition, $\|\Sigma\|_2 = \max_{\|x\| = 1} \|\Sigma x\|_2 = \max\{|\sigma_i|\} = \sigma_1$. Likewise, $\|A\|_F = \|\Sigma\|_F$, and by definition $\|\Sigma\|_F = \sqrt{\sigma_1^2 + \cdots + \sigma_r^2}$
Eigenvalue decomposition

- From linear algebra, $A\mathbf{x} = \lambda \mathbf{x}$, $\lambda$ is an eigenvalue, and $\mathbf{x}$ is an eigenvector.
- For $m$ eigenvectors,

$$A[x_1, x_2, \ldots, x_m] = [x_1, x_2, \ldots, x_m] \begin{bmatrix} \lambda_1 & & \\ & \lambda_2 & \\ & & \ddots \\ & & & \lambda_m \end{bmatrix}$$

and

$$AX = X\Lambda$$

where $\Lambda$ is an $m \times m$ diagonal matrix whose entries are the eigenvalues of $A$, and $X \in \mathbb{R}^{m \times m}$ contains linearly independent eigenvector of $A$.

- The eigenvalue decomposition of $A$

$$A = X\Lambda X^{-1}$$
SVD and eigenvalue decomposition

- SVD uses two different bases (the sets of left and right singular vectors), whereas the eigenvalue decomposition uses just one (eigenvectors)
- SVD uses orthonormal bases, whereas the eigenvalue decomposition uses a basis that generically is not orthogonal
- Not all matrices have an eigenvalue decomposition, but all matrices have a SVD
Matrix properties via SVD (cont’d)

**Theorem**

*The nonzero singular values of A are the square roots of the nonzero eigenvalues of $AA^\top$ or $A^\top A$ (they have the same nonzero eigenvalues).*

**Proof.**

From definition,

$$AA^\top = (U\Sigma V^\top)(U\Sigma V^\top)^\top = U\Sigma V^\top V\Sigma U^\top = U \text{ diag}(\sigma_1^2, \ldots, \sigma_p^2) \ U^\top$$

**Theorem**

*For $A \in \mathbb{R}^{m\times m}$, $|\det(A)| = \prod_{i=1}^m \sigma_i$*

**Proof.**

$$|\det(A)| = |\det(U\Sigma V^\top)| = |\det(U)||\det(\Sigma)||\det(V^\top)| = |\det(\Sigma)| = \prod_{i=1}^m \sigma_i$$
Low-rank approximation

**Theorem**

*(Eckart-Young 1936)* Let \( A = U \Sigma V^\top = U \text{ diag}(\sigma_1, \ldots, \sigma_r, 0, \ldots, 0) V^\top. \)

For any \( \nu \) with \( 0 \leq \nu \leq r \), \( A_\nu = \sum_{i=1}^{\nu} \sigma_i u_i v_i^\top \),

\[
\|A - A_\nu\|_2 = \min_{\text{rank}(B) \leq \nu} \|A - B\|_2 = \sigma_{\nu+1}
\]

**Proof.**

Suppose there is some \( B \) with \( \text{rank}(B) \leq \nu \) such that

\[
\|A - B\|_2 < \|A - A_\nu\|_2 = \sigma_{\nu+1}.
\]

Then there exists an \((n - \nu)\)-dimensional subspace \( W \in \mathbb{R}^n \) such that \( w \in W \Rightarrow Bw = 0 \). Then

\[
\|A w\|_2 = \|(A - B)w\|_2 \leq \|A - B\|_2 \|w\|_2 < \sigma_{\nu+1} \|w\|_2
\]

Thus \( W \) is a \((n - \nu)\)-dimensional subspace where \( \|A w\| < \sigma_{\nu+1} \|w\| \). But there is a \((\nu + 1)\)-dimensional subspace where \( \|A w\| \geq \sigma_{\nu+1} \|w\| \), namely the space spanned by the first \( \nu + 1 \) right singular vector of \( A \). Since the sum of the dimensions of these two spaces exceeds \( n \), there must be a nonzero vector lying in both, and this is a contradiction.
## Low-rank approximation

**Theorem**

A is the sum of r rank one matrices: \( A = \sum_{i=1}^{r} \sigma_i u_i v_j^\top \)

**Theorem (Eckart-Young 1936)** Let \( A = U \Sigma V^\top = U \text{ diag}(\sigma_1, \ldots, \sigma_r, 0, \ldots, 0) V^\top \). For any \( \nu \) with \( 0 \leq \nu \leq r \), \( A_\nu = \sum_{i=1}^{\nu} \sigma_i u_i v_i^\top \),

\[
\| A - A_\nu \|_2 = \min_{\text{rank}(B) \leq \nu} \| A - B \|_2 = \sigma_{\nu+1}
\]

**Proof.**

Let \( \Sigma_\nu = U (A - A_\nu) V^\top \), then

\[
\begin{align*}
\Sigma_\nu &= U \left( \text{diag}(\sigma_1, \ldots, \sigma_\nu, \sigma_{\nu+1}, \ldots, \sigma_p) - \text{diag}(\sigma_1, \ldots, \sigma_\nu, 0, \ldots, 0) \right) V^\top \\
&= U \text{ diag}(0, \ldots, 0, \sigma_{\nu+1}, \ldots, \sigma_p) V^\top
\end{align*}
\]

consequently \( \| A - A_\nu \|_2 = \| \Sigma_\nu \|_2 = \sigma_{\nu+1} \).
Geometric interpretation of Eckart-Young theorem

- What is the best approximation of a hyperellipsoid by a line segment?
  - Take the line segment to be the longest axis
- Next, what is the best approximation by a two-dimensional ellipsoid?
  - Take the ellipsoid spanned by the longest and the second longest axis
- Continue and improve the approximation by adding into our approximation the largest axis of the hyperellipsoid not yet included
- Reminiscent of techniques used in image compression, machine learning, and functional analysis (e.g., matching pursuit)

**Theorem**

For any $\nu$ with $0 \leq \nu \leq r$, $A_\nu = \sum_{i=1}^{\nu} \sigma_i u_i v_i^\top$, 

$$
\|A - A_\nu\|_F = \min_{\text{rank}(B) \leq \nu} \sqrt{\sigma_{\nu+1}^2 + \cdots + \sigma_r^2}
$$
Sensitivity of square systems

If
\[ A = \sum_{i=1}^{n} \sigma_i u_i v_i^\top = U \Sigma V^\top \]
is the SVD of A, then
\[ x = A^{-1} b = (U \Sigma V^\top)^{-1} b = \sum_{i=1}^{n} \frac{u_i^\top b}{\sigma_i} v_i \]

Small changes in A or b can induce relatively large changes in x if σₙ is small.

The magnitude of σₙ has bearing on the sensitivity of the Ax = b problem.

The solution x is increasingly sensitive to perturbations.
Condition

- Consider the parameterized system

\[(A + \varepsilon F)x(\varepsilon) = b + \varepsilon f \quad x(0) = x\]

where \(F \in \mathbb{R}^{n \times n}\) and \(f \in \mathbb{R}^n\)

- If \(A\) is nonsingular, then \(x(\varepsilon)\) is differentiable in a neighborhood of zero

- Moreover, \(\dot{x} = A^{-1}(f - Fx)\) and the Taylor series expansion

\[x(\varepsilon) = x + \varepsilon \dot{x}(0) + O(\varepsilon^2)\]

- Using any vector norm

\[
\frac{\|x(\varepsilon) - x\|}{\|x\|} \leq |\varepsilon| \|A^{-1}\| \left\{ \frac{\|f\|}{\|x\|} + \|F\| \right\} + O(\varepsilon^2)
\]
**Condition number**

- For square matrices $A$, define the condition number by
  \[ \kappa(A) = \|A\| \|A^{-1}\| \]
  
  with the convention that $\kappa(A) = \infty$ for singular $A$

- Using the inequality $\|b\| \leq \|A\| \|x\|$ it follows that
  \[ \frac{\|x(\varepsilon) - x\|}{\|x\|} \leq \kappa(A)(\rho_A + \rho_b) + O(\varepsilon^2) \]

  where
  \[ \rho_A = |\varepsilon| \frac{\|F\|}{\|A\|} \quad \text{and} \quad \rho_b = |\varepsilon| \frac{\|f\|}{\|b\|} \]

  represent the relative errors in $A$ and $b$

- The relative error in $x$ is $\kappa(A)$ times the relative error in $A$ and $b$

- The condition $\kappa(A)$ quantifies the sensitivity of the $Ax = b$ problem
Condition number (cont’d)

- Note that $\kappa(\cdot)$ depends on the underlying norm
  
  $$\kappa_2(A) = \|A\|_2 \|A^{-1}\|_2 = \frac{\sigma_1(A)}{\sigma_n(A)}$$

- $\kappa_2(A)$ measures the elongation of the hyperellipsoid \{\(Ax : \|x\|_2 = 1\}\}

- If $\kappa(A)$ is large, then $A$ is said to be an ill-conditioned matrix

- $\kappa_\alpha(\cdot)$ and $\kappa_\beta(\cdot)$ on $\mathbb{R}^{n \times n}$ are equivalent if constants $c_1$ and $c_2$ can be found such that $c_1 \kappa_\alpha(A) \leq \kappa_\beta(A) \leq c_2 \kappa_\alpha(A)$, e.g.,

  - $\frac{1}{n} \kappa_2(A) \leq \kappa_1(A) \leq n \kappa_2(A)$
  - $\frac{1}{n} \kappa_\infty(A) \leq \kappa_2(A) \leq n \kappa_\infty(A)$
  - $\frac{1}{n^2} \kappa_1(A) \leq \kappa_\infty(A) \leq n^2 \kappa_1(A)$

- For any $p$-norm, we have $\kappa(A) \geq 1$, and matrices with small conditional number are said to be well-conditioned.
Theorem

The minimum norm least squares solution to a linear system \( Ax = b \), that is, the shortest vector \( x \) that achieves \( \min_x \| Ax - b \| \) is unique, and is given by

\[
\hat{x} = V \Sigma^\dagger U^\top b
\]

where

\[
\Sigma^\dagger = \begin{bmatrix}
1/\sigma_1 & 0 & \cdots & 0 \\
& \ddots & & \\
& & 1/\sigma_r & \cdots \\
& & & 0 \\
& & & & \ddots \\
& & & & & 0 & 0 & \cdots & 0
\end{bmatrix}
\]

- The matrix \( A^\dagger = V \Sigma^\dagger U^\top \) is the pseudoinverse of \( A \)
The minimum norm solution to $Ax = b$ is the vector that minimizes $\|Ax - b\|$, 

$$\|U\Sigma V^\top x - b\| = \|U(\Sigma V^\top x - U^\top b)\|$$

and so it is equivalent to solve $\|\Sigma V^\top x - U^\top b\|$

Let $y = V^\top x$ and $c = U^\top b$, it becomes

$$\|\Sigma y - c\|$$

$$
\begin{bmatrix}
\sigma_1 & 0 & \cdots & 0 \\
0 & \ddots & \ddots & 0 \\
\vdots & & \ddots & 0 \\
0 & \cdots & & 0 \\
\end{bmatrix}
\begin{bmatrix}
y_1 \\
y_r \\
y_{r+1} \\
y_n \\
\end{bmatrix}
= 
\begin{bmatrix}
c_1 \\
c_r \\
c_{r+1} \\
c_m \\
\end{bmatrix}
$$
Minimum norm solution (cont’d)

- The optimal $y$ has the following components
  \[
  y_i = \frac{c_i}{\sigma_i} \quad \text{for } i = 1, \ldots, r \\
  y_i = 0 \quad \text{for } i = r + 1, \ldots, n
  \]

- In vector form
  \[
  y = \Sigma^\dagger c
  \]

- Notice there is no other choice for $y$, which is therefore unique:
  minimum residual forces the choice of $y_1, \ldots, y_r$, and minimum norm solution forces the other entries of $y$

- The minimum norm least squares solution is
  \[
  \hat{x} = V y = V \Sigma^\dagger c = V \Sigma^\dagger U^\top b
  \]

- The residual is
  \[
  \|A x - b\|^2 = \|\Sigma y - c\|^2 = \sum_{i=r+1}^{m} c_i^2 = \sum_{i=r+1}^{m} (u_i^\top b)^2
  \]

  which is the projection of $b$ onto the complement of the range of $A$
Least squares solution of homogeneous linear systems

**Theorem**

For $Ax = 0$ or $\min_{\|x\|=1} \|Ax\|$. Let $A = U\Sigma V^\top$, the solution is

$$x = \alpha_1 v_{n-k+1} + \ldots + \alpha_k v_n$$

where $k$ is the largest integer such that

$$\sigma_{n-k+1} = \ldots = \sigma_n,$$

and

$$\alpha_1^2 + \ldots + \alpha_k^2 = 1$$

**Proof.**

Consider the unit-norm least square solution

$$\|Ax\| = \|U\Sigma V^\top x\| \equiv \|\Sigma V^\top x\| = \|\Sigma y\|$$

where $y = V^\top x$. Thus the unit norm vector $y$ that minimizes the norm

$$\sigma_1^2 y_1^2 + \ldots + \sigma_n^2 y_n^2$$

which is achieved by concentrating all the mass of $y$ w.r.t smallest $\sigma$

$$y_1 = \ldots = y_{n-k} = 0$$

and thus $x = V y = y_1 v_1 + \ldots + y_{n-k+1} v_{n-k+1} + \ldots + y_n v_n$ and

$$\alpha_1 = y_{n-k+1}, \ldots, \alpha_k = y_n$$