Data-Driven Modeling and Control

Lecture 2

Dr. Shima Nazari

From Lecture 1

Recall pseudoinverse calculations for underdetermined systems:

$$Ax = b, \qquad \mathrm{A} \in \mathbb{R}^{m \times n}, x \in \mathbb{R}^{n}, b \in \mathbb{R}^{m}, \quad m < n$$

We seek to minimize norm solution from the infinitely many solutions:

$$\min \frac{1}{2} x^T x$$

$$s.t. \quad Ax - b = 0$$

Hamiltonian: $H = \frac{1}{2} x^T x + \lambda^T (Ax - b)$

From Lecture 1

Hamiltonian: $H = \frac{1}{2}x^T x + \lambda^T(Ax - b)$

$$\frac{\partial H}{\partial x} =$$

$$\frac{\partial H}{\partial \lambda} =$$

Rank and Nullity

Function:

A function $f: X \to Y$ assigns an element of Y to each element of X (called $f(x) = y$)

- Domain:
- Range

Rank and Nullity

Linear function (**linear operator**, linear transformation, linear mapping)

Operates on a linear space (V, F) to produce elements from another linear space (U, F).

Notated: $\mathcal{A}: (V, F) \to (U, F)$

Note:

Rank and Nullity

DEF: **Null space of A** is the set $\mathcal{N}(A) = \{v \in V, Av = \emptyset_u\}$

DEF: **Range of A** is the set $\mathcal{R}(A) = \{u \in U, Av = u \,\forall\, v \in V\}$

Rank and Nullity

Q: Can $\mathcal{N}(A)$ and $\mathcal{R}(A)$ be empty?

Rank and Nullity

DEF: The rank of the {linear map $\mathcal{A}$/ matrix A} is denoted by $rank(A)$ or $\rho(A)$ is the dimension of range of A.

$$\rho(A) =$$

DEF: The nullity of the {linear map $\mathcal{A}$/ matrix A} is denoted with $null(A)$ or $\nu(A)$ is the dimension of null space of A.

$$\nu(A) =$$

Rank and Nullity

Linearly independent columns of A:

Rank and Nullity

Rank-Nullity Theorem:

$$\dim\big(\mathcal{R}(A)\big) + \dim\big(\mathcal{N}(A)\big) = \dim(V)$$

Proof:

Rank-Nullity Theorem:

Rank-Nullity Theorem:

Rank-Nullity Theorem:

Rank-Nullity Theorem:

Eigenvalues and Eigenvectors

Let A to be a linear map that operates on a linear space (V, F) . Suppose that there exists $\lambda \in F$ and $v \in V$ $(v \neq \emptyset)$ such that:

$$Av = \lambda v$$

- λ is an **eigenvalue** of A
- v is an **eigenvector** of A

❖ For any eigenvalue there exists an associated linear subspace of eigenvectors

❖ $\mathcal{N}(A - \lambda I)$ is called $\mathcal{N}_\lambda$ (eigenspace)

❖ $\dim(\mathcal{N}_\lambda)$ is called the geometric multiplicity of λ

Eigenvalues and Eigenvectors

How to find eigenvalues of A:

I. write $Av = \lambda v$

Eigenvalues and Eigenvectors

II. Form $\Phi(\lambda) = \det(A - \lambda I)$, a polynomial of degree n in λs

II. Solutions to $\Phi(\lambda) = 0$ are eigen values of A.

IV. For distinct eigenvalues form $Av_i = \lambda_i v_i$ and fine v_is (eigenvectors)

Eigenvalues and Eigenvectors

Example: find eigenvalues and eigenvectors for $A = \begin{bmatrix} 1 & -1 \\ 2 & -1 \end{bmatrix}$

Eigenvalues and Eigenvectors

Eigenvalues and Eigenvectors

Theorem: If eigenvalues λ_i are all distinct, then the associated eigenvectors v_i in the set $\{v_1, v_2, \dots, v_n\}$ form a linearly independent basis.

Representation of A in $\{v_1, v_2, \dots, v_n\}$

Eigenvalues and Eigenvectors

Q: what if there are repeated eigenvalues? (not all eigenvalues are distinct)

Assume A: $V \rightarrow V$, $(\dim(V) = n)$

- A always has n eigenvalues (some distinct, some repeated)
- Let $\lambda_1, \lambda_2, \dots \lambda_c$ to be distinct eigenvalues of A
- Let $m_1, m_2, \dots m_c$ be algebraic multiplicity of λ_is.

Use $q_i = \dim\big(\mathcal{N}(A - \lambda_i I)\big)$

Eigenvalues and Eigenvectors

Case 1: $\boldsymbol{q_i = m_i}$

Eigenvalues and Eigenvectors

Case 1: $\boldsymbol{q_i < m_i}$ **or** $\boldsymbol{dim\big(\mathcal{N}(A - \lambda_i I)\big) < m_i}$

CLAIM 1: $\dim(\mathcal{N}(A - \lambda_i I)^{m_i}) = m_i$

Eigenvalues and Eigenvectors

DEF: A vector v is a **generalized eigenvector** of **rank k** associated with eigenvalue λ if:

$$\boldsymbol{v} \in \mathcal{N}(\boldsymbol{A} - \boldsymbol{\lambda I})^{\boldsymbol{k}} \text{ but } \boldsymbol{v} \notin \mathcal{N}(\boldsymbol{A} - \boldsymbol{\lambda I})^{\boldsymbol{k-1}}$$

Eigenvalues and Eigenvectors

CLAIM 2: $v^{\eta_i - 1} = (A - \lambda I)v^{\eta_i}$ is a generalized eigenvector of rank $\eta_i - 1$

Similarly:

- $\{v^1, v^2, \dots, v^{\eta_i}\}$ is called a chain of generalized eigenvectors.

Eigenvalues and Eigenvectors

How to construct a chain of eigenvectors:

Eigenvalues and Eigenvectors

CLAIM 3: Any chain of generalized eigenvectors are linearly independent.

CLAIM 4: chains associated with different eigenvalues are linearly independent and different chains associated with the same eigenvalue are linearly independent.

Eigenvalues and Eigenvectors

❖ Conclusion: set $\{v_1^1, \ldots, v_1^{m_1}, v_2^1, \ldots, v_2^{m_2}, \ldots\}$ are linearly independent basis.

What is the representation of A in about basis:

Eigenvalues and Eigenvectors

Singular Value Decomposition (SVD)

Adjoint of a linear operator $\mathrm{A}: V \rightarrow W$ is $A^*: W \rightarrow V$, such that $\forall x \in V$, and $\forall\, y \in W$:

$$< y, Ax > = < A^* y, x >$$

If $A \in \mathbb{C}^{n \times n}$, and we have $A = A^*$, then we say A is **Hermitian**.

If $A \in \mathbb{R}^{n \times n}$, and we have $A = A^T$, then we say A is **Symmetric**.

Singular Value Decomposition (SVD)

Theorem: If matrix A is Hermitian then all of its eigenvalues are real and it will have full set of eigenvectors that are all mutually orthogonal, regardless of the existence of any repeated eigenvalues.

Theorem: Hermitian matrices have a complete set of n regular (not generalized) eigenvectors.

Singular Value Decomposition (SVD)

Consider A^*A, where $A \in \mathbb{C}^{m \times n}$

CLAIM: The eigenvalues of A^*A are non-negative.

Singular Value Decomposition (SVD)

- It can be shown that: $\mathcal{R}(A^*) = \mathcal{R}(A^* A)$

$$\mathcal{N}(A) = \mathcal{N}(A^* A)$$

Singular Value Decomposition (SVD)

Consider $A \in \mathbb{C}^{m\times n}$ and $A^*A \in \mathbb{C}^{n\times n}$, assume that $\rho(A^*A) = r$, and let's name the eigenvalues of A^*A as $\sigma_1^2 \geq \sigma_2^2 \geq \cdots \geq \sigma_r^2 > \sigma_{r+1} = \cdots = \sigma_n = 0$

Singular Value Decomposition (SVD)

Representation of A^*A in the basis $\{v_i\}_1^n$:

Define $V_1 = [v_1, v_2, \dots, v_r]$ and $V_2 = [v_{r+1}, \dots, v_n]$

Singular Value Decomposition (SVD)

Define transformation matrix P

$$P^{-1} = [V_1 \mid V_2] = Q$$

Then

$$A^*A = Q \begin{bmatrix} \sigma^2 & 0 \\ 0 & 0 \end{bmatrix} Q^{-1}$$

Singular Value Decomposition (SVD)

We would like to show that there exists an orthonormal matrix $U \in \mathbb{C}^{m\times m}$ such that:

$$A = U\Sigma V^* = U\begin{bmatrix}\sigma & 0\\ 0 & 0\end{bmatrix}V^*$$

Review: $A^*A = V\begin{bmatrix}\sigma^2 & 0\\ 0 & 0\end{bmatrix}V^*$

Singular Value Decomposition (SVD)

Decompose $U = [U_1 | U_2]$ where $U_1 \in \mathbb{C}^{m\times r}$ and $U_2 \in \mathbb{C}^{m\times(m-r)}$

Then $A = U\Sigma V^* =$

Singular Value Decomposition (SVD)

- Columns of V_1 span $\mathcal{R}(A^*) = \mathcal{R}(A^*A)$, they are also eigenvectors A^*A corresponding to nonzero eigenvalues $\{\sigma_1^2, \sigma_2^2, \dots, \sigma_r^2\}$
- Columns of V_2 span $\mathcal{N}(A) = \mathcal{N}(A^*A)$, they are also eigenvectors A^*A corresponding to zero eigenvalues $\{\sigma_{r+1}^2, \dots, \sigma_n^2\} = \{0, \dots, 0\}$
- The columns of U_1 span $\mathcal{R}(A)$
- The columns of U_2 span $\mathcal{N}(A)$
- Singular values of A of are $\sigma_1, \dots, \sigma_n$, where $\sigma_i(A) = [\lambda_i(A^*A)]^{1/2}$, $i = 1, \dots, n$

Singular Value Decomposition (SVD)

Matrix Approximation

Theorem (Eckart-Young): The optimal rank-r approximation to matrix X, in a least squares sense, is given by the rank-p SVD truncation $\tilde{X}$:

$$\underset{\tilde{X}, s.t.\ rank(\tilde{X})=p}{\operatorname{argmin}} \left\| X - \tilde{X} \right\|_F = \tilde{U}\tilde{\Sigma}\tilde{V}^*$$

Where : $\tilde{V}$ and $\tilde{U}$ are the first p leading columns of V and U, and $\tilde{\Sigma}$ contains the leading $p \times p$ sub-block of Σ.

Review: Frobenius norm $\|.\|_F$

Singular Value Decomposition (SVD)

$\tilde{X}$ is an approximation to X and is called truncated SVD of X:

$$\tilde{X} = \tilde{U}\tilde{\Sigma}\tilde{V}^*$$

Singular Value Decomposition (SVD)

Applications of SVD:

- Image compression.
- Market data analysis.
- Latent Semantic Indexing (LSI).
- Political spectrum analysis.
- 3D image deformation using moving least-squares.
- SVD and PCA for gene expression data.

See example applications in chapter 1 of your textbook and here:
https://inst.eecs.berkeley.edu/~ee127/sp21/livebook/l_svd_apps.html

Singular Value Decomposition (SVD)

Example: Image compression

Let $X \in \mathbb{R}^{n\times m}$ be a real valued matrix representing pixel values for the image. n and m are the number of pixels in the vertical and horizontal directions.

Singular Value Decomposition (SVD)

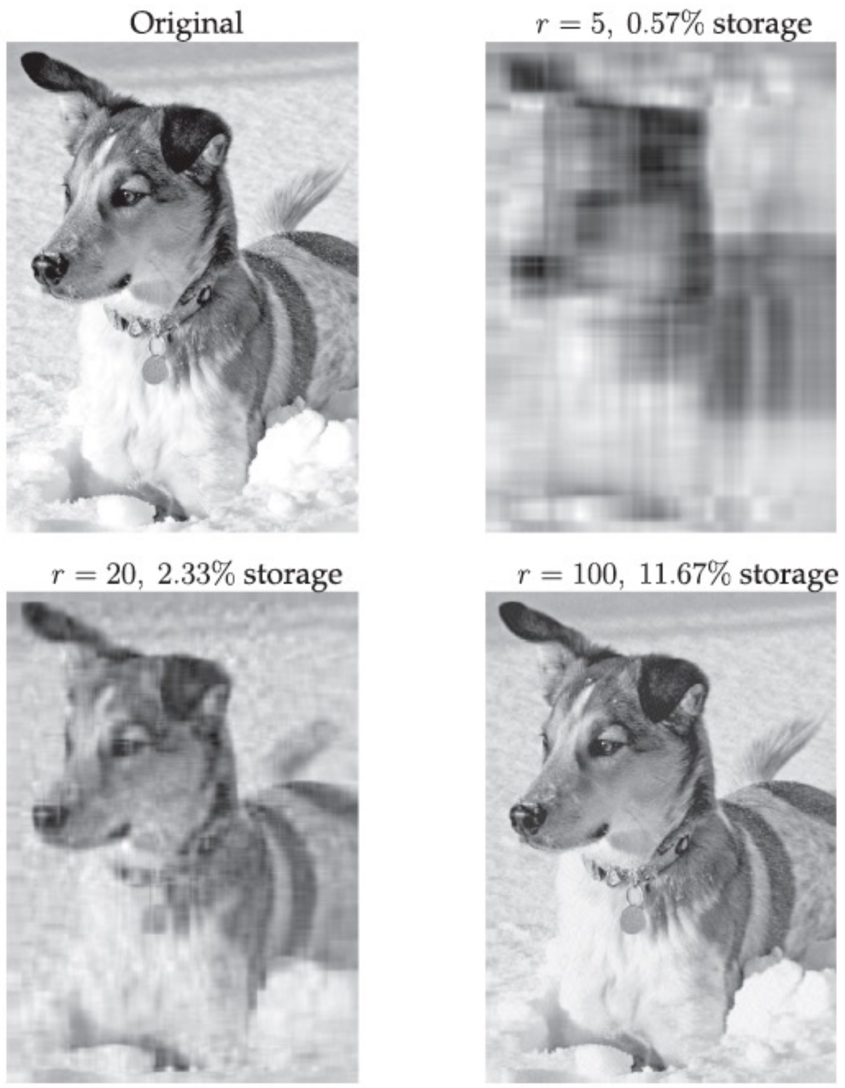

Figure 1.3 Image compression of Mordecai the snow dog, truncating the SVD at various ranks r. Original image resolution is 2000×1500.