E9 211: Adaptive Signal Processing

Lecture 2: Linear Algebra -II

Outline

1. QR factorization
2. Singular value decomposition
3. Connection between eigenvalue decomposition, QR, and SVD
4. Pseudo-inverse

QR factorization

- Let $\mathbf{A}$ be an $N \times N$ square full rank matrix.
  Then there is a decomposition

$$\mathbf{A} = [\mathbf{a}_1 \ \mathbf{a}_2 \cdots \mathbf{a}_N] = [\mathbf{q}_1 \ \mathbf{q}_2 \cdots \mathbf{q}_N] \begin{bmatrix} r_{11} & r_{12} & \cdots & r_{1N} \\ 0 & r_{22} & \cdots & r_{2N} \\ 0 & 0 & \ddots & \vdots \\ 0 & 0 & 0 & r_{NN} \end{bmatrix}$$

  Here, $\mathbf{Q}$ is a unitary matrix, $\mathbf{R}$ is upper triangular and square.

- Interpretation:
  - $\mathbf{q}_1$ is a normalized vector with the same direction as $\mathbf{a}_1$.
  - $[\mathbf{q}_1 \ \mathbf{q}_2]$ is an isometry spanning the same space as $[\mathbf{a}_1 \ \mathbf{a}_2]$.
  - So on.

QR decomposition

- Let $\mathbf{A}$ be an $M \times N$ tall $(M \geq N)$ matrix.
  Then there is a decomposition

$$\mathbf{A} = \mathbf{QR} = [\hat{\mathbf{Q}}\ \hat{\mathbf{Q}}^{\perp}] \begin{bmatrix} \hat{\mathbf{R}} \\ \mathbf{0} \end{bmatrix} = \hat{\mathbf{Q}}\hat{\mathbf{R}}$$

  Here, $\mathbf{Q}$ is a unitary matrix, $\hat{\mathbf{R}}$ is upper triangular and square.

- Properties:
  - $\mathbf{R}$ is upper triangular with $M - N$ zeros added.
  - $\mathbf{A} = \hat{\mathbf{Q}}\hat{\mathbf{R}}$ is a "reduced" or an "economy-sized" QR decomposition
  - If $\hat{\mathbf{R}}$ is full rank, the columns of $\hat{\mathbf{Q}}$ span the range of $\mathbf{A}$.

Singular value decomposition

- For any matrix $\mathbf{X}$, there is a decomposition

$$\mathbf{X} = \mathbf{U}\boldsymbol{\Sigma}\mathbf{V}^{\mathrm{H}}$$

  Here, $\mathbf{U}$ and $\mathbf{V}$ are unitary, and $\boldsymbol{\Sigma}$ is diagonal with positive real entries.

- Properties:
  - The columns $\mathbf{u}_i$ of $\mathbf{U}$ are called the left singular vectors
  - The columns $\mathbf{v}_i$ of $\mathbf{V}$ are called the right singular vectors
  - The diagonal entries σ_i of $\boldsymbol{\Sigma}$ are called the singular values
  - They are positive, real, and sorted

$$\sigma_1 \geq \sigma_2 \geq \cdots 0$$

Singular value decomposition

- For an $M \times N$ tall matrix $\mathbf{X}$, there is a decomposition

$$\mathbf{X} = \mathbf{U}\boldsymbol{\Sigma}\mathbf{V}^{\mathrm{H}} = [\hat{\mathbf{U}}\ \hat{\mathbf{U}}^{\perp}] \left[\begin{array}{cc|cc} \sigma_1 & & & \\ & \sigma_d & & \\ \hline & & 0 & \\ & & & 0 \\ \hline 0 & \cdots & \cdots & 0 \\ 0 & \cdots & \cdots & 0 \end{array}\right] \left[\begin{array}{c} \hat{\mathbf{V}} \\ (\hat{\mathbf{V}}^{\perp})^{\mathrm{H}} \end{array}\right]$$

$$\mathbf{U} : M \times M, \quad \boldsymbol{\Sigma} : M \times N, \mathbf{V} : N \times N$$

$$\sigma_1 \geq \sigma_2 \geq \cdots \sigma_d > \sigma_{d+1} = \cdots \sigma_N = 0$$

- Economy size SVD: $\mathbf{X} = \hat{\mathbf{U}}\hat{\boldsymbol{\Sigma}}\hat{\mathbf{V}}^{\mathrm{H}}$, where $\hat{\boldsymbol{\Sigma}} : d \times d$ is a diagonal matrix containing $\sigma_1, \cdots, \sigma_d$ along the diagonals.

Singular value decomposition

- The rank of $\mathbf{X}$ is d, the number of nonzero singular values
- $\mathbf{X} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^{\mathrm{H}} \;\Leftrightarrow\; \mathbf{X}^{\mathrm{H}} = \mathbf{V}\mathbf{\Sigma}\mathbf{U}^{\mathrm{H}} \;\Leftrightarrow\; \mathbf{X}\mathbf{V} = \mathbf{U}\mathbf{\Sigma} \;\Leftrightarrow\; \mathbf{X}^{\mathrm{H}}\mathbf{U} = \mathbf{V}\mathbf{\Sigma}$
  - The columns of $\hat{\mathbf{U}}$ ($\hat{\mathbf{U}}^{\perp}$) are the orthonormal basis for $\mathrm{ran}(\mathbf{X})$ ($\ker(\mathbf{X}^{\mathrm{H}})$)
  - The columns of $\hat{\mathbf{V}}$ ($\hat{\mathbf{V}}^{\perp}$) are the orthonormal basis for $\mathrm{ran}(\mathbf{X}^{\mathrm{H}})$ ($\ker(\mathbf{X})$)
- $\mathbf{X} = \sum_{i=1}^{d} \sigma_i(\mathbf{u}_i\mathbf{v}_i^{\mathrm{H}})$; $\mathbf{u}_i\mathbf{v}_i^{\mathrm{H}}$ is a rank-1 isometry matrix.
- $\mathbf{X}\mathbf{v}_i = \sigma_i\mathbf{u}_i$
- $\|\mathbf{X}\| = \|\mathbf{X}^{\mathrm{H}}\| = \sigma_1$, the largest singular value.

QR and SVD

- The QR factorization of a tall $(M \geq N)$ matrix $\mathbf{X}$ is

$$\mathbf{X} = \mathbf{QR} = [\hat{\mathbf{Q}}\ \hat{\mathbf{Q}}^{\perp}] \begin{bmatrix} \hat{\mathbf{R}} \\ \mathbf{0} \end{bmatrix} = \hat{\mathbf{Q}}\hat{\mathbf{R}}$$

- Compute SVD of $\hat{\mathbf{R}}$: $M \times M$

$$\hat{\mathbf{R}} = \hat{\mathbf{U}}_R \hat{\mathbf{\Sigma}}_R \hat{\mathbf{V}}_R^{\mathrm{H}}$$

- The SVD of $\mathbf{X}$ is

$$\mathbf{X} = (\hat{\mathbf{Q}}\hat{\mathbf{U}}_R)\hat{\mathbf{\Sigma}}_R \hat{\mathbf{V}}_R^{\mathrm{H}}$$

  Obviously, $\mathbf{X}$ and $\mathbf{R}$ have the same left singular vectors and singular values.

Eigenvalue decomposition

- The eigenvalue problem is $(\mathbf{A} - \lambda \mathbf{I})\mathbf{x} = \mathbf{0}$
- Any λ that makes $\mathbf{A} - \lambda \mathbf{I}$ singular is called an eigenvalue and the corresponding invariant vector is called the eigenvector
- Stacking

$$\mathbf{A}[\mathbf{x}_1 \ \mathbf{x}_2 \cdots] = [\mathbf{x}_1 \ \mathbf{x}_2 \cdots] \begin{bmatrix} \lambda_1 & & \\ & \lambda_2 & \\ & & \ddots \end{bmatrix}$$

$$\mathbf{AT} = \mathbf{T\Lambda} \Leftrightarrow \mathbf{A} = \mathbf{T\Lambda T}^{-1}$$

(might exist when $\mathbf{T}$ is invertible and when eigenvalues are distinct)

Eigenvalue decomposition and QR

- Suppose $\mathbf{T}$ has QR factorization $\mathbf{T} = \mathbf{Q}\mathbf{R}_T$ or $\mathbf{T}^{-1} = \mathbf{R}_T^{-1}\mathbf{Q}^H$. Therefore

  $$\mathbf{A} = \mathbf{T}\mathbf{\Lambda}\mathbf{T}^{-1} = \mathbf{Q}\mathbf{R}_T\mathbf{\Lambda}\mathbf{R}_T^{-1}\mathbf{Q}^H = \mathbf{Q}\mathbf{R}\mathbf{Q}^{\mathrm{H}}$$: **Schur decomposition**

  $\mathbf{Q}$ is unitary and $\mathbf{R}$ is upper triangular.

  - $\mathbf{R}$ has eigenvalues of $\mathbf{A}$ along the diagonal
  - Schur decomposition always exists.
  - $\mathbf{Q}$ does not contain the eigenvectors, but has information about the eigen-subspaces.

Eigenvalue decomposition and SVD

- Suppose the SVD of $\mathbf{X} = \mathbf{U}\boldsymbol{\Sigma}\mathbf{V}^{\mathrm{H}}$. Therefore

$$\mathbf{X}\mathbf{X}^{\mathrm{H}} = \mathbf{U}\boldsymbol{\Sigma}\mathbf{V}^{\mathrm{H}}\mathbf{V}^{\mathrm{H}}\boldsymbol{\Sigma}\mathbf{U}^{\mathrm{H}} = \mathbf{U}\boldsymbol{\Sigma}^{2}\mathbf{U}^{\mathrm{H}} = \mathbf{U}\boldsymbol{\Lambda}\mathbf{U}^{\mathrm{H}}$$

  - The eigenvalues of $\mathbf{X}\mathbf{X}^{\mathrm{H}}$ are singular values of $\mathbf{X}$ squared.
  - Eigenvectors of $\mathbf{X}\mathbf{X}^{\mathrm{H}}$ are the left singular vectors of $\mathbf{X}$
  - Eigenvalue decomposition of $\mathbf{X}\mathbf{X}^{\mathrm{H}}$ always exits and SVD always exists.

Pseudo inverse

- For a tall full-column rank matrix $\mathbf{X} : M \times N$

  Pseudo-inverse of $\mathbf{X}$ is $\mathbf{X}^{\dagger} = (\mathbf{X}^{\mathrm{H}}\mathbf{X})^{-1}\mathbf{X}^{\mathrm{H}}$.
  - $\mathbf{X}^{\dagger}\mathbf{X} = \mathbf{I}_N$:inverse on the short space)
  - $\mathbf{X}\mathbf{X}^{\dagger} = \mathbf{P}_c$: Projector onto $\mathrm{ran}(\mathbf{X})$
- For a tall rank matrix $\mathbf{X} : M \times N$ with rank d, $\mathbf{X}\mathbf{X}^{\mathrm{H}}$ is not invertible.

  Moore-Penrose Pseudo inverse of $\mathbf{X} = \hat{\mathbf{U}}\hat{\boldsymbol{\Sigma}}\hat{\mathbf{V}}$ is $\mathbf{X}^{\dagger} = \hat{\mathbf{V}}\hat{\boldsymbol{\Sigma}}^{-1}\hat{\mathbf{U}}^{\mathrm{H}}$
  1. $\mathbf{X}\mathbf{X}^{\dagger}\mathbf{X} = \mathbf{X}$
  2. $\mathbf{X}^{\dagger}\mathbf{X}\mathbf{X}^{\dagger} = \mathbf{X}^{\dagger}$
  3. $\mathbf{X}\mathbf{X}^{\dagger} = \hat{\mathbf{U}}\hat{\mathbf{U}}^{\mathrm{H}} = \mathbf{P}_c$:Projector onto $\mathrm{ran}(\mathbf{X})$
  4. $\mathbf{X}^{\dagger}\mathbf{X}^{\dagger} = \hat{\mathbf{V}}\hat{\mathbf{V}}^{\mathrm{H}} = \mathbf{P}_r$:Projector onto $\mathrm{ran}(\mathbf{X}^{\mathrm{H}})$