Understanding EVD: Change of Basis

syn. to an invertible matrix.
In part., columns of X are linearly independent.

Let $X \in \mathbb{C}^{n\times n}$ be a nonsingular matrix.

- The columns $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n$ of X form a basis of $\mathbb{C}^n$.
- Any $\mathbf{z} \in \mathbb{C}^n$ is uniquely written as

X nonsingular

$X^{-1}\vec{z} = \vec{u}$

$$\mathbf{z} = X\mathbf{u} = u_1\mathbf{x}_1 + u_2\mathbf{x}_2 + \cdots + u_n\mathbf{x}_n.$$

(linear comb. of basis vectors)

$$= \begin{bmatrix} \vec{x}_1 & \vec{x}_2 & \cdots & \vec{x}_n \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{bmatrix}$$

  - The scalars $u_1, \ldots, u_n$ are called the **coordinates** of $\mathbf{z}$ with respect to the columns of X.
  - The vector $\mathbf{u} = X^{-1}\mathbf{z}$ is the representation of $\mathbf{z}$ with respect to the basis consisting of the columns of X.

Upshot

Left-multiplication by X^{-1} performs a **change of basis** into the coordinates associated with the columns of X.

Understanding EVD: Change of Basis (cont')

V-basis = basis consisting of columns of V $= \{\vec{v}_1, \vec{v}_2, \cdots, \vec{v}_n\}$

Suppose $A \in \mathbb{C}^{n\times n}$ has an EVD $A = VDV^{-1}$. Then, for any $\mathbf{z} \in \mathbb{C}^n$, $\mathbf{y} = A\mathbf{z}$ can be written as

(i.e., A is diagonalizable)

$$\boxed{\underbrace{V^{-1}\mathbf{y}}_{\substack{\text{coord. of } \vec{y} \\ \text{w.r.t. V-basis}}} = D\underbrace{V^{-1}\mathbf{z}}_{\substack{\text{coord. of } \vec{z} \\ \text{w.r.t. V-basis}}}.}$$

$\vec{y} = A\vec{z}$

$= VDV^{-1}\vec{z}$

$V^{-1}\vec{y} = V^{-1}VDV^{-1}\vec{z}$

$V^{-1}V = I$

Interpretation

The matrix A is a diagonal transformation in the coordinates with respect to the V-basis.

basis consisting of eigenvectors

* diagonal transformation : coordinates are rescaled independently

For example

$$D = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 5 & 0 \\ 0 & 0 & 2 \end{bmatrix}$$

$D : \vec{e}_1 \longmapsto \vec{e}_1 \quad (D\vec{e}_1 = \vec{e}_1)$

$D : \vec{e}_2 \longmapsto 5\vec{e}_2$

$D : \vec{e}_3 \longmapsto 2\vec{e}_3$

What is EVD good for?

`[V, D] = eig(A);`

Q. Why do we like diagonalizable matrices?

$$A = VDV^{-1}, \quad V \text{ nonsingular}, \quad D \text{ diagonal}.$$

Situation: Matrix powers.

- $$A^2 = (VDV^{-1})(VDV^{-1})$$
$$= V D \underbrace{(V^{-1}V)}_{=I} D V^{-1}$$
$$= V D^2 V^{-1}$$

- $$A^3 = (VD\underbrace{V^{-1})(V}_{=I}D\underbrace{V^{-1})(V}_{=I}DV^{-1})$$
$$= V D^3 V^{-1}$$

- In general,
$$A^k = V D^k V^{-1}$$

$$D = \begin{bmatrix} \lambda_1 & & & 0 \\ & \lambda_2 & & \\ & & \ddots & \\ 0 & & & \lambda_n \end{bmatrix}$$

$$D^k = \begin{bmatrix} \lambda_1^k & & & \\ & \lambda_2^k & & \\ & & \ddots & \\ & & & \lambda_n^k \end{bmatrix}$$

Lec 36: Spectral Theory

Singular Value Decomposition (SVD)

Singular Value Decomposition: Overview

Singular Value Decomposition

$A = LU$, $A = QR$, $A = VDV^{-1}$

Theorem 1 (SVD)

Let $A \in \mathbb{C}^{m\times n}$. Then A can be written as

$$A = U\Sigma V^*, \tag{SVD}$$

where $U \in \mathbb{C}^{m\times m}$ and $V \in \mathbb{C}^{n\times n}$ are unitary and $\Sigma \in \mathbb{R}^{m\times n}$ is diagonal. If A is real, then so are U and V.

- The columns of U are called the **left singular vectors** of A;
- The columns of V are called the **right singular vectors** of A;
- The diagonal entries of Σ, written as $\sigma_1, \sigma_2, \ldots, \sigma_r$, for $r = \min\{m, n\}$, are called the **singular values** of A and they are nonnegative numbers ordered as

$$\sigma_1 \geqslant \sigma_2 \geqslant \cdots \geqslant \sigma_r \geqslant 0.$$

$A = VDV^{-1}$

SVD (real)

$A \in \mathbb{R}^{m\times n}$, $m > n$ (more rows than cols ; tall rect.)

$$\boxed{A = U \Sigma V^T}$$

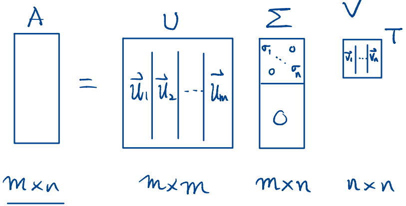
where

- $U \in \mathbb{R}^{m\times m}$ orthogonal ($U^T U = I \in \mathbb{R}^{m\times m}$)
- $\Sigma \in \mathbb{R}^{m\times n}$ diagonal
- $V \in \mathbb{R}^{n\times n}$ orthogonal ($V^T V = I \in \mathbb{R}^{n\times n}$)

Orthogonal matrices $\longrightarrow$ QR.

Cartoon View

$V^{-1} = V^T$



Terms

- $\vec{u}_1, \cdots, \vec{u}_m$: left singular vectors
- $\vec{v}_1, \cdots, \vec{v}_n$: right singular vectors
- $\sigma_1, \cdots, \sigma_n$: singular values

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

Generalization to Complex SVD

$A \in \mathbb{C}^{m \times n}$, $m > n$

$$\boxed{A = U \Sigma V^*}$$

where

- $U \in \mathbb{C}^{m \times m}$ unitary
- $\Sigma \in \mathbb{R}^{m \times n}$ diagonal
- $V \in \mathbb{C}^{n \times n}$ unitary

Note 1

$V^* = (\bar{V})^T$

- Conjugate transpose of V
- a.k.a. the hermitian of V.

Note 2

A complex $U \in \mathbb{C}^{m \times m}$ is unitary if $U^* U = I$.

(columns of U are orthonormal.)

$\vec{u}^* \vec{v}$ inner product in complex vectors

Thick vs Thin SVD

Suppose that $m > n$ and observe that:

$$A = U \Sigma V^* \quad \text{(thick SVD)}$$

Note that $V^* V = I$,

$$AV = U\Sigma$$

$$= \left[\, \underbrace{\vec{u}_1 \cdots \vec{u}_n}_{\hat{U}} \;\middle|\; \underbrace{\vec{u}_{n+1} \cdots \vec{u}_m}_{\hat{U}^{\perp}} \,\right] \begin{bmatrix} \overset{\hat{\Sigma}}{\ddots} \\ \hline 0 \end{bmatrix} \begin{matrix} \}\ n \text{ row} \\ \}\ 0 \end{matrix} = \hat{U} \hat{\Sigma}$$

Thus,

$$A = \hat{U} \hat{\Sigma} V^* \quad \text{(thin SVD)}.$$

SVD in MATLAB

- Thick SVD: `[U,S,V] = svd(A);`
- Thin SVD: `[U,S,V] = svd(A, 0);`

Understanding SVD

Geometric Perspective

Write $A = U\Sigma V^*$ as $AV = U\Sigma$:

$$A\mathbf{v}_k = \sigma_k \mathbf{u}_k, \quad k = 1, \ldots, r = \min\{m, n\}.$$

The image of the unit sphere under any $m \times n$ matrix is a hyperellipse.

Algebraic Perspective

Alternately, note that $\mathbf{y} = A\mathbf{z} \in \mathbb{C}^m$ for any $\mathbf{z} \in \mathbb{C}^n$ can be written as

$$(U^* \mathbf{y}) = \Sigma \left(V^* \mathbf{z}\right).$$

Any matrix $A \in \mathbb{C}^{m \times n}$ can be viewed as a diagonal transformation from $\mathbb{C}^n$ (source space) to $\mathbb{C}^m$ (target space) with respect to suitably chosen orthonormal bases for both spaces.

SVD vs. EVD

Recall that a diagonalizable $A = VDV^{-1} \in \mathbb{C}^{n\times n}$ satisfies

$$\mathbf{y} = A\mathbf{z} \quad \longrightarrow \quad \left(V^{-1}\mathbf{y}\right) = D\left(V^{-1}\mathbf{z}\right).$$

This allowed us to view any diagonalizable square matrix $A \in \mathbb{C}^{n\times n}$ as a diagonal transformation from $\mathbb{C}^n$ to itself[1] with respect to the basis formed by a set of eigenvector of A.

Differences.

- **Basis:** SVD uses two ONBs (left and right singular vectors); EVD uses one, usually non-orthogonal basis (eigenvectors).
- **Universality:** all matrices have an SVD; not all matrices have an EVD.
- **Utility:** SVD is useful in problems involving the behavior of A or A^+; EVD is relevant to problems involving A^k.

[1] The source and the target spaces of the transformation coincide.