LINEAR ALGEBRA -II

B V Rajarama Bhat

Indian Statistical Institute, Bangalore

Lecture 29: Singular value decomposition

- First we recall the notion of positivity and some characterizations.

Lecture 29: Singular value decomposition

- First we recall the notion of positivity and some characterizations.
- Definition 26.1: An $n \times n$ matrix A is said to be a positive matrix if $A = B^*B$ for some $n \times n$ matrix B.

- First we recall the notion of positivity and some characterizations.
- Definition 26.1: An $n \times n$ matrix A is said to be a positive matrix if $A = B^*B$ for some $n \times n$ matrix B.
- Some authors may call these as non-negative definite matrices and invertible matrices of the form B^*B as positive definite matrices.

- First we recall the notion of positivity and some characterizations.
- Definition 26.1: An $n \times n$ matrix A is said to be a positive matrix if $A = B^*B$ for some $n \times n$ matrix B.
- Some authors may call these as non-negative definite matrices and invertible matrices of the form B^*B as positive definite matrices.
- Warning: A positive matrix need not have positive entries. It can have negative entries and also complex entries.

Lecture 29: Singular value decomposition

- First we recall the notion of positivity and some characterizations.
- Definition 26.1: An $n \times n$ matrix A is said to be a positive matrix if $A = B^*B$ for some $n \times n$ matrix B.
- Some authors may call these as non-negative definite matrices and invertible matrices of the form B^*B as positive definite matrices.
- Warning: A positive matrix need not have positive entries. It can have negative entries and also complex entries.
- Matrices whose entries are positive would be called as entrywise positive matrices. That is also an important class, but we will not be studying them now.

Characterizations of positivity

- Theorem 26.2: Let $A = [a_{ij}]_{1 \leq i,j \leq n}$ be a complex matrix. Then the following are equivalent:

- Theorem 26.2: Let $A = [a_{ij}]_{1 \leq i,j \leq n}$ be a complex matrix. Then the following are equivalent:
- (i) A is positive, that is, $A = B^*B$ for some $n \times n$ matrix B.

Characterizations of positivity

- Theorem 26.2: Let $A = [a_{ij}]_{1 \leq i,j \leq n}$ be a complex matrix. Then the following are equivalent:
- (i) A is positive, that is, $A = B^*B$ for some $n \times n$ matrix B.
- (ii) $A = C^*C$ for some $m \times n$ matrix for some m.

Characterizations of positivity

- Theorem 26.2: Let $A = [a_{ij}]_{1 \leq i,j \leq n}$ be a complex matrix. Then the following are equivalent:
- (i) A is positive, that is, $A = B^*B$ for some $n \times n$ matrix B.
- (ii) $A = C^*C$ for some $m \times n$ matrix for some m.
- (iii) $a_{ij} = \langle v_i, v_j \rangle, 1 \leq i,j \leq n$ for vectors $v_1, v_2, \ldots, v_n$ in some inner product space V.

Characterizations of positivity

- Theorem 26.2: Let $A = [a_{ij}]_{1 \leq i,j \leq n}$ be a complex matrix. Then the following are equivalent:
- (i) A is positive, that is, $A = B^*B$ for some $n \times n$ matrix B.
- (ii) $A = C^*C$ for some $m \times n$ matrix for some m.
- (iii) $a_{ij} = \langle v_i, v_j \rangle, 1 \leq i, j \leq n$ for vectors $v_1, v_2, \ldots, v_n$ in some inner product space V.
- (iv) $\langle x, Ax \rangle \geq 0$ for all $x \in \mathbb{C}^n$.

Characterizations of positivity

- Theorem 26.2: Let $A = [a_{ij}]_{1 \leq i,j \leq n}$ be a complex matrix. Then the following are equivalent:
- (i) A is positive, that is, $A = B^*B$ for some $n \times n$ matrix B.
- (ii) $A = C^*C$ for some $m \times n$ matrix for some m.
- (iii) $a_{ij} = \langle v_i, v_j \rangle, 1 \leq i, j \leq n$ for vectors $v_1, v_2, \ldots, v_n$ in some inner product space V.
- (iv) $\langle x, Ax \rangle \geq 0$ for all $x \in \mathbb{C}^n$.
- (v) $A = A^*$ and eigenvalues of A are non-negative.

Characterizations of positivity

- Theorem 26.2: Let $A = [a_{ij}]_{1 \leq i,j \leq n}$ be a complex matrix. Then the following are equivalent:
- (i) A is positive, that is, $A = B^*B$ for some $n \times n$ matrix B.
- (ii) $A = C^*C$ for some $m \times n$ matrix for some m.
- (iii) $a_{ij} = \langle v_i, v_j \rangle, 1 \leq i, j \leq n$ for vectors $v_1, v_2, \ldots, v_n$ in some inner product space V.
- (iv) $\langle x, Ax \rangle \geq 0$ for all $x \in \mathbb{C}^n$.
- (v) $A = A^*$ and eigenvalues of A are non-negative.
- (vi) $A = S^2$ for some self-adjoint $n \times n$ matrix S.

Continuation

- Proof: (i) $\Rightarrow$ (ii). Take $m = n$ and $C = B$.

Continuation

- Proof: (i) $\Rightarrow$ (ii). Take $m = n$ and $C = B$.
- (ii) $\Rightarrow$ (iii). Let $v_1, v_2, \ldots, v_n$ be the columns of C. Then $v_j \in \mathbb{C}^m$ for every j and $A = C^*C$ implies $a_{ij} = \langle v_i, v_j \rangle$.

Continuation

- Proof: (i) $\Rightarrow$ (ii). Take $m = n$ and $C = B$.
- (ii) $\Rightarrow$ (iii). Let $v_1, v_2, \ldots, v_n$ be the columns of C. Then $v_j \in \mathbb{C}^m$ for every j and $A = C^*C$ implies $a_{ij} = \langle v_i, v_j \rangle$.
- (iii) $\Rightarrow$ (iv). We have $a_{ij} = \langle v_i, v_j \rangle, \quad \forall i, j$.

Continuation

- Proof: (i) $\Rightarrow$ (ii). Take $m = n$ and $C = B$.
- (ii) $\Rightarrow$ (iii). Let $v_1, v_2, \ldots, v_n$ be the columns of C. Then $v_j \in \mathbb{C}^m$ for every j and $A = C^*C$ implies $a_{ij} = \langle v_i, v_j \rangle$.
- (iii) $\Rightarrow$ (iv). We have $a_{ij} = \langle v_i, v_j \rangle, \quad \forall i, j$.
- Now for any $x \in \mathbb{C}^n$:

$$\begin{aligned} \langle x, Ax \rangle &= \sum_{i=1}^{n} \overline{x_i}(Ax)_i \\ &= \sum_{i=1}^{n} \overline{x_i}. \sum_{j=1}^{n} a_{ij} x_j \\ &= \sum_{i=1}^{n} \overline{x_i}. \sum_{j=1}^{n} \langle v_i, v_j \rangle . x_j \end{aligned}$$

Continuation

- Therefore,

$$\begin{aligned}\langle x, Ax\rangle &= \sum_{i=1}^{n}\sum_{j=1}^{n}\langle x_i v_i, x_j v_j\rangle \\ &= \langle\sum_{i=1}^{n} x_i v_i, \sum_{j=1}^{n} x_j v_j\rangle \\ &= \langle y, y\rangle\end{aligned}$$

Continuation

- Therefore,

$$\begin{aligned}\langle x, Ax\rangle &= \sum_{i=1}^{n}\sum_{j=1}^{n}\langle x_i v_i, x_j v_j\rangle \\ &= \langle\sum_{i=1}^{n} x_i v_i, \sum_{j=1}^{n} x_j v_j\rangle \\ &= \langle y, y\rangle\end{aligned}$$

- where $y = \sum_{i=1}^{n} x_i v_i$.

Continuation

- Therefore,

$$\begin{aligned} \langle x, Ax \rangle &= \sum_{i=1}^{n} \sum_{j=1}^{n} \langle x_i v_i, x_j v_j \rangle \\ &= \langle \sum_{i=1}^{n} x_i v_i, \sum_{j=1}^{n} x_j v_j \rangle \\ &= \langle y, y \rangle \end{aligned}$$

- where $y = \sum_{i=1}^{n} x_i v_i$.
- Hence

$$\langle x, Ax \rangle \geq 0.$$

Continuation

- (iv) $\Rightarrow$ (v). It is given that $\langle x, Ax \rangle \geq 0$ for every $x \in \mathbb{C}^n$.

Continuation

- (iv) $\Rightarrow$ (v). It is given that $\langle x, Ax\rangle \geq 0$ for every $x \in \mathbb{C}^n$.
- First we want to show that $A = A^*$. Here we use the polarization identity and the fact that if $\langle v, w\rangle$ is real then $\langle v, w\rangle = \langle w, v\rangle$. For all x, y,

$$\begin{aligned}
\langle x, Ay\rangle &= \frac{1}{4}\sum_{j=0}^{3} i^{-j}\langle (x+i^j y), A(x+i^j y)\rangle \\
&= \frac{1}{4}\sum_{j=0}^{3} i^{-j}\langle A(x+i^j y), (x+i^j y)\rangle \\
&= \langle Ax, y\rangle
\end{aligned}$$

Continuation

- (iv) $\Rightarrow$ (v). It is given that $\langle x, Ax\rangle \geq 0$ for every $x \in \mathbb{C}^n$.
- First we want to show that $A = A^*$. Here we use the polarization identity and the fact that if $\langle v, w\rangle$ is real then $\langle v, w\rangle = \langle w, v\rangle$. For all x, y,

$$\begin{aligned} \langle x, Ay\rangle &= \frac{1}{4}\sum_{j=0}^{3} i^{-j}\langle (x + i^j y), A(x + i^j y)\rangle \\ &= \frac{1}{4}\sum_{j=0}^{3} i^{-j}\langle A(x + i^j y), (x + i^j y)\rangle \\ &= \langle Ax, y\rangle \end{aligned}$$

- This proves $A^* = A$ from the defining condition of the adjoint.

Continuation

- (iv) $\Rightarrow$ (v). It is given that $\langle x, Ax\rangle \geq 0$ for every $x \in \mathbb{C}^n$.
- First we want to show that $A = A^*$. Here we use the polarization identity and the fact that if $\langle v, w\rangle$ is real then $\langle v, w\rangle = \langle w, v\rangle$. For all x, y,

$$\begin{aligned}
\langle x, Ay\rangle &= \frac{1}{4}\sum_{j=0}^{3} i^{-j}\langle (x + i^j y), A(x + i^j y)\rangle \\
&= \frac{1}{4}\sum_{j=0}^{3} i^{-j}\langle A(x + i^j y), (x + i^j y)\rangle \\
&= \langle Ax, y\rangle
\end{aligned}$$

- This proves $A^* = A$ from the defining condition of the adjoint.
- Now suppose a is an eigenvalue of A. Choose an eigenvector x with a as the eigenvalue. Then

$$\langle x, Ax\rangle = a\langle x, x\rangle \geq 0.$$

Continuation

- (iv) $\Rightarrow$ (v). It is given that $\langle x, Ax\rangle \geq 0$ for every $x \in \mathbb{C}^n$.
- First we want to show that $A = A^*$. Here we use the polarization identity and the fact that if $\langle v, w\rangle$ is real then $\langle v, w\rangle = \langle w, v\rangle$. For all x, y,

$$\begin{aligned}\langle x, Ay\rangle &= \frac{1}{4}\sum_{j=0}^{3} i^{-j}\langle (x + i^j y), A(x + i^j y)\rangle \\ &= \frac{1}{4}\sum_{j=0}^{3} i^{-j}\langle A(x + i^j y), (x + i^j y)\rangle \\ &= \langle Ax, y\rangle\end{aligned}$$

- This proves $A^* = A$ from the defining condition of the adjoint.
- Now suppose a is an eigenvalue of A. Choose an eigenvector x with a as the eigenvalue. Then

$$\langle x, Ax\rangle = a\langle x, x\rangle \geq 0.$$

- implies that $a \geq 0$ as $\langle x, x\rangle \neq 0$.

Continuation

- (v) $\Rightarrow$ (vi). We assume $A = A^*$ and the eigenvalues of A are non-negative.

Continuation

- (v) $\Rightarrow$ (vi). We assume $A = A^*$ and the eigenvalues of A are non-negative.
- By spectral theorem there exists a unitary U and a diagonal matrix D, such that
$$A = UDU^*.$$

Continuation

- (v) $\Rightarrow$ (vi). We assume $A = A^*$ and the eigenvalues of A are non-negative.
- By spectral theorem there exists a unitary U and a diagonal matrix D, such that
$$A = UDU^*.$$
- Since the eigenvalues of A are non-negative, the diagonal entries of D are non-negative. We denote the diagonal entries by $d_1, d_2, \ldots, d_n$.

Continuation

- (v) $\Rightarrow$ (vi). We assume $A = A^*$ and the eigenvalues of A are non-negative.
- By spectral theorem there exists a unitary U and a diagonal matrix D, such that
$$A = UDU^*.$$
- Since the eigenvalues of A are non-negative, the diagonal entries of D are non-negative. We denote the diagonal entries by $d_1, d_2, \ldots, d_n$.
- Take
$$S = U \begin{bmatrix} \sqrt{d_1} & 0 & \dots & 0 \\ 0 & \sqrt{d_2} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \sqrt{d_n} \end{bmatrix} U^*.$$

Continuation

- (v) $\Rightarrow$ (vi). We assume $A = A^*$ and the eigenvalues of A are non-negative.
- By spectral theorem there exists a unitary U and a diagonal matrix D, such that
$$A = UDU^*.$$
- Since the eigenvalues of A are non-negative, the diagonal entries of D are non-negative. We denote the diagonal entries by $d_1, d_2, \ldots, d_n$.
- Take
$$S = U \begin{bmatrix} \sqrt{d_1} & 0 & \ldots & 0 \\ 0 & \sqrt{d_2} & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & \sqrt{d_n} \end{bmatrix} U^*.$$
- Then clearly S is self-adjoint and $A = S^2$.

Continuation

- (vi) $\Rightarrow$ (i). Assume $A = S^2$ where $S = S^*$.

Continuation

- (vi) $\Rightarrow$ (i). Assume $A = S^2$ where $S = S^*$.
- In particular, $A = S^*S$. This proves (i). $\blacksquare$

Continuation

- (vi) $\Rightarrow$ (i). Assume $A = S^2$ where $S = S^*$.
- In particular, $A = S^*S$. This proves (i). ∎
- Example 26.3: Take

$$R = \begin{bmatrix} 2 & -1 \\ -1 & 2 \end{bmatrix}.$$

Continuation

- (vi) $\Rightarrow$ (i). Assume $A = S^2$ where $S = S^*$.
- In particular, $A = S^*S$. This proves (i). $\blacksquare$
- Example 26.3: Take

$$R = \begin{bmatrix} 2 & -1 \\ -1 & 2 \end{bmatrix}.$$

- Clearly R is self-adjoint. We have the characteristic polynomial of R, as

$$p(x) = (x-2)^2 - 1 = x^2 - 4x + 3 = (x-1)(x-3).$$

Continuation

- (vi) $\Rightarrow$ (i). Assume $A = S^2$ where $S = S^*$.
- In particular, $A = S^* S$. This proves (i). ■
- Example 26.3: Take

$$R = \begin{bmatrix} 2 & -1 \\ -1 & 2 \end{bmatrix}.$$

- Clearly R is self-adjoint. We have the characteristic polynomial of R, as

$$p(x) = (x-2)^2 - 1 = x^2 - 4x + 3 = (x-1)(x-3).$$

- Therefore, the eigenvalues of R are $\{1, 3\}$, which are non-negative. Then by part (v) of the previous Theorem, R is positive.

Continuation

- (vi) $\Rightarrow$ (i). Assume $A = S^2$ where $S = S^*$.
- In particular, $A = S^*S$. This proves (i). ■
- Example 26.3: Take

$$R = \begin{bmatrix} 2 & -1 \\ -1 & 2 \end{bmatrix}.$$

- Clearly R is self-adjoint. We have the characteristic polynomial of R, as

$$p(x) = (x-2)^2 - 1 = x^2 - 4x + 3 = (x-1)(x-3).$$

- Therefore, the eigenvalues of R are $\{1, 3\}$, which are non-negative. Then by part (v) of the previous Theorem, R is positive.
- Note that though R is positive as per our definition, some of its entries are negative.

Continuation

- (vi) $\Rightarrow$ (i). Assume $A = S^2$ where $S = S^*$.
- In particular, $A = S^*S$. This proves (i). ■
- Example 26.3: Take

$$R = \begin{bmatrix} 2 & -1 \\ -1 & 2 \end{bmatrix}.$$

- Clearly R is self-adjoint. We have the characteristic polynomial of R, as

$$p(x) = (x-2)^2 - 1 = x^2 - 4x + 3 = (x-1)(x-3).$$

- Therefore, the eigenvalues of R are $\{1, 3\}$, which are non-negative. Then by part (v) of the previous Theorem, R is positive.
- Note that though R is positive as per our definition, some of its entries are negative.
- Find all self-adjoint operators S such that $R = S^2$. (Exercise)

Cartesian decomposition

- Theorem 26.3 (Cartesian decomposition): Let A be a complex square matrix. Then A decomposes uniquely as

$$A = B + iC$$

where B, C are self-adjoint.

Cartesian decomposition

- Theorem 26.3 (Cartesian decomposition): Let A be a complex square matrix. Then A decomposes uniquely as

$$A = B + iC$$

where B, C are self-adjoint.
- Proof: Take $B = \frac{A+A^*}{2}$ and $C = \frac{A-A^*}{2i}$.

Cartesian decomposition

- Theorem 26.3 (Cartesian decomposition): Let A be a complex square matrix. Then A decomposes uniquely as

$$A = B + iC$$

where B, C are self-adjoint.
- Proof: Take $B = \frac{A+A^*}{2}$ and $C = \frac{A-A^*}{2i}$.
- Then it is easily verified that B, C are self-adjoint and $A = B + iC$.

Cartesian decomposition

- Theorem 26.3 (Cartesian decomposition): Let A be a complex square matrix. Then A decomposes uniquely as

$$A = B + iC$$

  where B, C are self-adjoint.
- Proof: Take $B = \frac{A+A^*}{2}$ and $C = \frac{A-A^*}{2i}$.
- Then it is easily verified that B, C are self-adjoint and $A = B + iC$.
- Conversely, suppose $A = B + iC$, with B, C self-adjoint. We see directly that $B = \frac{A+A^*}{2}$ and $C = \frac{A-A^*}{2i}$.

Cartesian decomposition

- Theorem 26.3 (Cartesian decomposition): Let A be a complex square matrix. Then A decomposes uniquely as

$$A = B + iC$$

where B, C are self-adjoint.

- Proof: Take $B = \frac{A+A^*}{2}$ and $C = \frac{A-A^*}{2i}$.
- Then it is easily verified that B, C are self-adjoint and $A = B + iC$.
- Conversely, suppose $A = B + iC$, with B, C self-adjoint. We see directly that $B = \frac{A+A^*}{2}$ and $C = \frac{A-A^*}{2i}$.
- This proves uniqueness. ■

Cartesian decomposition

- Theorem 26.3 (Cartesian decomposition): Let A be a complex square matrix. Then A decomposes uniquely as

$$A = B + iC$$

where B, C are self-adjoint.
- Proof: Take $B = \frac{A+A^*}{2}$ and $C = \frac{A-A^*}{2i}$.
- Then it is easily verified that B, C are self-adjoint and $A = B + iC$.
- Conversely, suppose $A = B + iC$, with B, C self-adjoint. We see directly that $B = \frac{A+A^*}{2}$ and $C = \frac{A-A^*}{2i}$.
- This proves uniqueness. ■
- This is known as Cartesian decomposition.

Constructing positive matrices out of positive matrices

- Theorem 27.1 Suppose A, B are positive matrices and a, b are positive real numbers. Then $aA + bB$ is positive.

Constructing positive matrices out of positive matrices

- Theorem 27.1 Suppose A, B are positive matrices and a, b are positive real numbers. Then $aA + bB$ is positive.
- Proof: Suppose A, B are $n \times n$ matrices. Consider arbitrary $x \in \mathbb{C}^n$.

Constructing positive matrices out of positive matrices

- Theorem 27.1 Suppose A, B are positive matrices and a, b are positive real numbers. Then $aA + bB$ is positive.
- Proof: Suppose A, B are $n \times n$ matrices. Consider arbitrary $x \in \mathbb{C}^n$.
- Then, $\langle x, (aA + bB)x \rangle = a\langle x, Ax \rangle + b\langle x, Bx \rangle \geq 0$

Constructing positive matrices out of positive matrices

- Theorem 27.1 Suppose A, B are positive matrices and a, b are positive real numbers. Then $aA + bB$ is positive.
- Proof: Suppose A, B are $n \times n$ matrices. Consider arbitrary $x \in \mathbb{C}^n$.
- Then, $\langle x, (aA + bB)x \rangle = a\langle x, Ax \rangle + b\langle x, Bx \rangle \geq 0$
- as the quadratic forms $\langle x, Ax \rangle$, and $\langle x, Bx \rangle$ are positive.

Constructing positive matrices out of positive matrices

- Theorem 27.1 Suppose A, B are positive matrices and a, b are positive real numbers. Then $aA + bB$ is positive.
- Proof: Suppose A, B are $n \times n$ matrices. Consider arbitrary $x \in \mathbb{C}^n$.
- Then, $\langle x, (aA + bB)x \rangle = a\langle x, Ax \rangle + b\langle x, Bx \rangle \geq 0$
- as the quadratic forms $\langle x, Ax \rangle$, and $\langle x, Bx \rangle$ are positive.
- Then by part (iv) of the characterization theorem, $aA + bB$ is positive. $\blacksquare$

Constructing positive matrices out of positive matrices

- Theorem 27.1 Suppose A, B are positive matrices and a, b are positive real numbers. Then $aA + bB$ is positive.
- Proof: Suppose A, B are $n \times n$ matrices. Consider arbitrary $x \in \mathbb{C}^n$.
- Then, $\langle x, (aA + bB)x \rangle = a\langle x, Ax \rangle + b\langle x, Bx \rangle \geq 0$
- as the quadratic forms $\langle x, Ax \rangle$, and $\langle x, Bx \rangle$ are positive.
- Then by part (iv) of the characterization theorem, $aA + bB$ is positive. ■
- Note that this theorem does not follow directly from the definition of positivity or from the eigenvalue criterion.

Constructing positive matrices out of positive matrices

- Theorem 27.1 Suppose A, B are positive matrices and a, b are positive real numbers. Then $aA + bB$ is positive.
- Proof: Suppose A, B are $n \times n$ matrices. Consider arbitrary $x \in \mathbb{C}^n$.
- Then, $\langle x, (aA + bB)x \rangle = a\langle x, Ax \rangle + b\langle x, Bx \rangle \geq 0$
- as the quadratic forms $\langle x, Ax \rangle$, and $\langle x, Bx \rangle$ are positive.
- Then by part (iv) of the characterization theorem, $aA + bB$ is positive. $\blacksquare$
- Note that this theorem does not follow directly from the definition of positivity or from the eigenvalue criterion.
- This theorem shows that the set of $n \times n$ positive matrices has 'cone' structure: It is closed under taking sums and it is closed under multiplication by positive scalar.

Products

- Remark 27.2: Suppose A, B are $n \times n$ positive matrices. Then AB need not be positive as it may not be self-adjoint.

Products

- Remark 27.2: Suppose A, B are $n \times n$ positive matrices. Then AB need not be positive as it may not be self-adjoint.
- Theorem 27.3: Let A be an $n \times n$ positive matrix. Suppose B is an $n \times m$ matrix. Then B^*AB is positive.

Products

- Remark 27.2: Suppose A, B are $n \times n$ positive matrices. Then AB need not be positive as it may not be self-adjoint.
- Theorem 27.3: Let A be an $n \times n$ positive matrix. Suppose B is an $n \times m$ matrix. Then B^*AB is positive.
- Proof: As A is positive, $A = D^*D$ for some matrix D.

Products

- Remark 27.2: Suppose A, B are $n \times n$ positive matrices. Then AB need not be positive as it may not be self-adjoint.
- Theorem 27.3: Let A be an $n \times n$ positive matrix. Suppose B is an $n \times m$ matrix. Then B^*AB is positive.
- Proof: As A is positive, $A = D^*D$ for some matrix D.
- Now, $B^*AB = B^*D^*DB = (DB)^*(DB)$. Hence B^*AB is positive from the definition of positivity. We may also see this from looking at the quadratic form. ■

Trace and Determinant

- Theorem 27.4: Let A be a positive matrix. Then trace $(A) \geq 0$ and $\det(A) \geq 0$. Diagonal entries of A are positive.

Trace and Determinant

- Theorem 27.4: Let A be a positive matrix. Then trace $(A) \geq 0$ and $\det(A) \geq 0$. Diagonal entries of A are positive.
- Proof: The first part is clear as the trace and determinant of a matrix are respectively the sum and the product of its eigenvalues and a positive matrix has non-negative eigenvalues. The second claim follows from $a_{ii} = \langle v_i, v_i \rangle$ in part (iv) of the characterization.

Gram matrices

- Definition 27.5: Let $v_1, v_2, \ldots, v_n$ be vectors in an inner product space V. Then their Gram matrix is defined as the matrix
$$A = [\langle v_i, v_j \rangle]_{1 \leq i,j \leq n}.$$

Gram matrices

- Definition 27.5: Let $v_1, v_2, \ldots, v_n$ be vectors in an inner product space V. Then their Gram matrix is defined as the matrix
$$A = [\langle v_i, v_j \rangle]_{1 \leq i,j \leq n}.$$
- We have seen that Gram matrices are positive and conversely all positive matrices can be written as Gram matrices. In Probability theory Gram matrices appear as 'covariance matrices'.

Gram matrices

- Definition 27.5: Let $v_1, v_2, \ldots, v_n$ be vectors in an inner product space V. Then their Gram matrix is defined as the matrix
$$A = [\langle v_i, v_j \rangle]_{1 \leq i,j \leq n}.$$
- We have seen that Gram matrices are positive and conversely all positive matrices can be written as Gram matrices. In Probability theory Gram matrices appear as 'covariance matrices'.
- Suppose x, y are vectors in an inner product space V. Consider their Gram matrix:
$$G = \begin{bmatrix} \langle x, x \rangle & \langle x, y \rangle \\ \langle y, x \rangle & \langle y, y \rangle \end{bmatrix}.$$

Gram matrices

- Definition 27.5: Let $v_1, v_2, \ldots, v_n$ be vectors in an inner product space V. Then their Gram matrix is defined as the matrix
$$A = [\langle v_i, v_j \rangle]_{1 \leq i,j \leq n}.$$
- We have seen that Gram matrices are positive and conversely all positive matrices can be written as Gram matrices. In Probability theory Gram matrices appear as 'covariance matrices'.
- Suppose x, y are vectors in an inner product space V. Consider their Gram matrix:
$$G = \begin{bmatrix} \langle x, x \rangle & \langle x, y \rangle \\ \langle y, x \rangle & \langle y, y \rangle \end{bmatrix}.$$
- We know that G is positive. Hence its determinant is positive. So we get $\langle x, x \rangle . \langle y, y \rangle - \langle x, y \rangle . \langle y, x \rangle \geq 0$.

Gram matrices

- Definition 27.5: Let $v_1, v_2, \ldots, v_n$ be vectors in an inner product space V. Then their Gram matrix is defined as the matrix
$$A = [\langle v_i, v_j \rangle]_{1 \leq i,j \leq n}.$$
- We have seen that Gram matrices are positive and conversely all positive matrices can be written as Gram matrices. In Probability theory Gram matrices appear as 'covariance matrices'.
- Suppose x, y are vectors in an inner product space V. Consider their Gram matrix:
$$G = \begin{bmatrix} \langle x, x \rangle & \langle x, y \rangle \\ \langle y, x \rangle & \langle y, y \rangle \end{bmatrix}.$$
- We know that G is positive. Hence its determinant is positive. So we get $\langle x, x \rangle . \langle y, y \rangle - \langle x, y \rangle . \langle y, x \rangle \geq 0$.
- In other words, we have the Cauchy-Schwarz inequality:
$$|\langle x, y \rangle|^2 \leq \|x\|^2 . \|y\|^2.$$

Uniqueness of positive square root

- Theorem 27.5: Let A be a positive matrix. Then there exists unique positive S such that $A = S^2$.

Uniqueness of positive square root

- Theorem 27.5: Let A be a positive matrix. Then there exists unique positive S such that $A = S^2$.
- Proof: In the proof the main characterization theorem for positive matrices we have seen that if A is positive then there exists positive S such that $A = S^2$.

Uniqueness of positive square root

- Theorem 27.5: Let A be a positive matrix. Then there exists unique positive S such that $A = S^2$.
- Proof: In the proof the main characterization theorem for positive matrices we have seen that if A is positive then there exists positive S such that $A = S^2$.
- Now suppose B is positive and $A = B^2$.

Uniqueness of positive square root

- Theorem 27.5: Let A be a positive matrix. Then there exists unique positive S such that $A = S^2$.
- Proof: In the proof the main characterization theorem for positive matrices we have seen that if A is positive then there exists positive S such that $A = S^2$.
- Now suppose B is positive and $A = B^2$.
- Let $b_1, b_2, \ldots, b_k$ the distinct eigenvalues of B and $B = b_1 Q_1 + b_2 Q_2 + \cdots + b_k Q_k$ be the spectral decomposition of B.

Continuation

- We have,

$$A = B^2 = b_1^2 Q_1 + b_2^2 Q_2 + \cdots + b_k^2 Q_k.$$

Continuation

- We have,

$$A = B^2 = b_1^2 Q_1 + b_2^2 Q_2 + \cdots + b_k^2 Q_k.$$

- Note that as $b_1, b_2, \ldots, b_k$ are positive and distinct, $b_1^2, b_2^2, \ldots, b_k^2$ are also positive and distinct.

Continuation

- We have,

$$A = B^2 = b_1^2 Q_1 + b_2^2 Q_2 + \cdots + b_k^2 Q_k.$$

- Note that as $b_1, b_2, \ldots, b_k$ are positive and distinct, $b_1^2, b_2^2, \ldots, b_k^2$ are also positive and distinct.
- Hence $A = b_1^2 Q_1 + b_2^2 Q_2 + \cdots + b_k^2 Q_k$ is the unique (up to permutation) spectral decomposition of A.

Continuation

- We have,

$$A = B^2 = b_1^2 Q_1 + b_2^2 Q_2 + \cdots + b_k^2 Q_k.$$

- Note that as $b_1, b_2, \ldots, b_k$ are positive and distinct, $b_1^2, b_2^2, \ldots, b_k^2$ are also positive and distinct.
- Hence $A = b_1^2 Q_1 + b_2^2 Q_2 + \cdots + b_k^2 Q_k$ is the unique (up to permutation) spectral decomposition of A.
- In other words, if $A = a_1 P_1 + a_2 P_2 + \cdots + a_k P_k$ is the spectral decomposition of A, then $B = \sqrt{a_1} P_1 + \sqrt{a_2} P_2 + \cdots + \sqrt{a_k} P_k$.

Continuation

- We have,

$$A = B^2 = b_1^2 Q_1 + b_2^2 Q_2 + \cdots + b_k^2 Q_k.$$

- Note that as $b_1, b_2, \ldots, b_k$ are positive and distinct, $b_1^2, b_2^2, \ldots, b_k^2$ are also positive and distinct.
- Hence $A = b_1^2 Q_1 + b_2^2 Q_2 + \cdots + b_k^2 Q_k$ is the unique (up to permutation) spectral decomposition of A.
- In other words, if $A = a_1 P_1 + a_2 P_2 + \cdots + a_k P_k$ is the spectral decomposition of A, then $B = \sqrt{a_1} P_1 + \sqrt{a_2} P_2 + \cdots + \sqrt{a_k} P_k$.
- This proves uniqueness. ■

Continuation

- Note that if we do not insist on positivity of the square root there can be many square roots.

Continuation

- Note that if we do not insist on positivity of the square root there can be many square roots.
- For instance, for A positive, if $A = UDU^*$ is the diagonalization of A,

$$U \begin{bmatrix} \pm\sqrt{d_1} & 0 & \cdots & 0 \\ 0 & \pm\sqrt{d_2} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \pm\sqrt{d_n} \end{bmatrix} U^*$$

with any choice of sign on the diagonal is a square root of A.

Continuation

- Note that if we do not insist on positivity of the square root there can be many square roots.
- For instance, for A positive, if $A = UDU^*$ is the diagonalization of A,

$$U \begin{bmatrix} \pm\sqrt{d_1} & 0 & \cdots & 0 \\ 0 & \pm\sqrt{d_2} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \pm\sqrt{d_n} \end{bmatrix} U^*$$

  with any choice of sign on the diagonal is a square root of A.
- For any projection P, the unitary $2P - I = P - P^{\perp}$ is a square root of I. This shows that I has infinitely many square roots (in dimension bigger than 1) if we do not insist on positivity.

Notation

- Notation: For a positive matrix A, $A^{\frac{1}{2}}$ denotes the square root of A. For any matrix B, $|B|$ denotes $(B^*B)^{\frac{1}{2}}$.

Notation

- Notation: For a positive matrix A, $A^{\frac{1}{2}}$ denotes the square root of A. For any matrix B, $|B|$ denotes $(B^*B)^{\frac{1}{2}}$.
- Note that for a positive matrix A, $A^{\frac{1}{2}} = A$ if and only if A is a projection.

Notation

- Notation: For a positive matrix A, $A^{\frac{1}{2}}$ denotes the square root of A. For any matrix B, $|B|$ denotes $(B^*B)^{\frac{1}{2}}$.
- Note that for a positive matrix A, $A^{\frac{1}{2}} = A$ if and only if A is a projection.
- Question: Do we have $|A+B| \leq |A| + |B|$? In other words can we say that $|A| + |B| - |A+B|$ is positive?

Polar decomposition theorem

- Theorem 28.1: Let B be an $n \times n$ matrix. Then B factorizes as
$$B = UR$$
where U is a unitary and R is positive. Here $R = |B|$ and U is uniquely determined if B is invertible.

Polar decomposition theorem

- Theorem 28.1: Let B be an $n \times n$ matrix. Then B factorizes as
$$B = UR$$
where U is a unitary and R is positive. Here $R = |B|$ and U is uniquely determined if B is invertible.
- Proof: We want to find U, R satisfying
$$B = UR$$
as above.

Polar decomposition theorem

- Theorem 28.1: Let B be an $n \times n$ matrix. Then B factorizes as
$$B = UR$$
where U is a unitary and R is positive. Here $R = |B|$ and U is uniquely determined if B is invertible.
- Proof: We want to find U, R satisfying
$$B = UR$$
as above.
- As suggested by the second line we take $R = |B| = (B^*B)^{\frac{1}{2}}$.

Polar decomposition theorem

- Theorem 28.1: Let B be an $n \times n$ matrix. Then B factorizes as
$$B = UR$$
where U is a unitary and R is positive. Here $R = |B|$ and U is uniquely determined if B is invertible.
- Proof: We want to find U, R satisfying
$$B = UR$$
as above.
- As suggested by the second line we take $R = |B| = (B^*B)^{\frac{1}{2}}$.
- Let $\mathcal{M}$ be the range of $|B|$, that is,
$$\mathcal{M} = \{|B|x : x \in \mathbb{C}^n\}.$$

Polar decomposition theorem

- Theorem 28.1: Let B be an $n \times n$ matrix. Then B factorizes as
$$B = UR$$
where U is a unitary and R is positive. Here $R = |B|$ and U is uniquely determined if B is invertible.
- Proof: We want to find U, R satisfying
$$B = UR$$
as above.
- As suggested by the second line we take $R = |B| = (B^*B)^{\frac{1}{2}}$.
- Let $\mathcal{M}$ be the range of $|B|$, that is,
$$\mathcal{M} = \{|B|x : x \in \mathbb{C}^n\}.$$
- Clearly $\mathcal{M}$ is a subspace of $\mathbb{C}^n$.

Continuation

- Similarly let $\mathcal{N}$ be the range of B:

$$\mathcal{N} = \{Bx : x \in \mathbb{C}^n\}.$$

Continuation

- Similarly let $\mathcal{N}$ be the range of B:

$$\mathcal{N} = \{Bx : x \in \mathbb{C}^n\}.$$

- Define $U_0 : \mathcal{M} \to \mathcal{N}$ by

$$U_0|B|x = Bx, x \in \mathbb{C}^n.$$

Continuation

- Similarly let $\mathcal{N}$ be the range of B:

$$\mathcal{N} = \{Bx : x \in \mathbb{C}^n\}.$$

- Define $U_0 : \mathcal{M} \to \mathcal{N}$ by

$$U_0|B|x = Bx, x \in \mathbb{C}^n.$$

- We claim that U_0 is a well-defined unitary from $\mathcal{M}$ to $\mathcal{N}$.

Continuation

- Similarly let $\mathcal{N}$ be the range of B:

$$\mathcal{N} = \{Bx : x \in \mathbb{C}^n\}.$$

- Define $U_0 : \mathcal{M} \to \mathcal{N}$ by

$$U_0|B|x = Bx, x \in \mathbb{C}^n.$$

- We claim that U_0 is a well-defined unitary from $\mathcal{M}$ to $\mathcal{N}$.
- Observe that, for $u, v \in \mathbb{C}^n$,

$$\begin{aligned} \langle Bu, Bv \rangle &= \langle u, B^* Bv \rangle \\ &= \langle u, |B|^2 v \rangle \\ &= \langle |B|u, |B|v \rangle \end{aligned}$$

Continuation

- Similarly let $\mathcal{N}$ be the range of B:
$$\mathcal{N} = \{Bx : x \in \mathbb{C}^n\}.$$
- Define $U_0 : \mathcal{M} \to \mathcal{N}$ by
$$U_0|B|x = Bx, x \in \mathbb{C}^n.$$
- We claim that U_0 is a well-defined unitary from $\mathcal{M}$ to $\mathcal{N}$.
- Observe that, for $u, v \in \mathbb{C}^n$,
$$\begin{aligned} \langle Bu, Bv\rangle &= \langle u, B^*Bv\rangle \\ &= \langle u, |B|^2 v\rangle \\ &= \langle |B|u, |B|v\rangle \end{aligned}$$
- In particular, if $|B|x = |B|y$, then $|B|(x - y) = 0$, Hence,
$$\langle |B|(x - y), |B|(x - y)\rangle = 0.$$

Continuation

- Taking $u = v = (x - y)$, in the previous equation we get $\langle B(x-y)B(x-y)\rangle = 0$ or $B(x-y) = 0$, that is, $Bx = By$.

Continuation

- Taking $u = v = (x - y)$, in the previous equation we get $\langle B(x-y)B(x-y)\rangle = 0$ or $B(x-y) = 0$, that is, $Bx = By$.
- In other words if $|B|x = |B|y$ then $Bx = By$. This shows that U_0 is well defined.

Continuation

- Taking $u = v = (x - y)$, in the previous equation we get $\langle B(x-y)B(x-y)\rangle = 0$ or $B(x-y) = 0$, that is, $Bx = By$.
- In other words if $|B|x = |B|y$ then $Bx = By$. This shows that U_0 is well defined.
- Clearly U_0 as defined is linear.

Continuation

- Taking $u = v = (x - y)$, in the previous equation we get $\langle B(x-y)B(x-y)\rangle = 0$ or $B(x-y) = 0$, that is, $Bx = By$.
- In other words if $|B|x = |B|y$ then $Bx = By$. This shows that U_0 is well defined.
- Clearly U_0 as defined is linear.
- Further,

$$\langle |B|u, |B|v\rangle = \langle Bu, Bv\rangle, \quad \forall u, v \in \mathbb{C}^n$$

shows that

$$\langle U_0(|B|u), U_0(|B|v)\rangle = \langle |B|u, |B|v\rangle, \quad \forall u, v \in \mathbb{C}^n$$

Continuation

- Taking $u = v = (x - y)$, in the previous equation we get $\langle B(x-y)B(x-y)\rangle = 0$ or $B(x-y) = 0$, that is, $Bx = By$.
- In other words if $|B|x = |B|y$ then $Bx = By$. This shows that U_0 is well defined.
- Clearly U_0 as defined is linear.
- Further,

$$\langle |B|u, |B|v\rangle = \langle Bu, Bv\rangle, \quad \forall u, v \in \mathbb{C}^n$$

  shows that

$$\langle U_0(|B|u), U_0(|B|v)\rangle = \langle |B|u, |B|v\rangle, \quad \forall u, v \in \mathbb{C}^n$$

- Hence U_0 is isometric.

Continuation

- Taking $u = v = (x - y)$, in the previous equation we get $\langle B(x-y)B(x-y)\rangle = 0$ or $B(x-y) = 0$, that is, $Bx = By$.
- In other words if $|B|x = |B|y$ then $Bx = By$. This shows that U_0 is well defined.
- Clearly U_0 as defined is linear.
- Further,

$$\langle |B|u, |B|v\rangle = \langle Bu, Bv\rangle, \quad \forall u, v \in \mathbb{C}^n$$

shows that

$$\langle U_0(|B|u), U_0(|B|v)\rangle = \langle |B|u, |B|v\rangle, \quad \forall u, v \in \mathbb{C}^n$$

- Hence U_0 is isometric.
- From the definition of U_0 it is clear that U_0 maps $\mathcal{M}$ onto $\mathcal{N}$.

Continuation

- We can extend U_0 to a unitary from $\mathbb{C}^n$ to $\mathbb{C}^n$ as the dimensions of $\mathcal{M}^\perp$ and $\mathcal{N}^\perp$ are equal. More explicitly we can take the following steps:

Continuation

- We can extend U_0 to a unitary from $\mathbb{C}^n$ to $\mathbb{C}^n$ as the dimensions of $\mathcal{M}^\perp$ and $\mathcal{N}^\perp$ are equal. More explicitly we can take the following steps:
- Let $v_1, v_2, \ldots, v_k$ be an ortho-normal basis of $\mathcal{M}$. Then $U_0 v_1, U_0 v_2, \ldots, U_0 v_k$ is an orthonormal basis of $\mathcal{N}$.

Continuation

- We can extend U_0 to a unitary from $\mathbb{C}^n$ to $\mathbb{C}^n$ as the dimensions of $\mathcal{M}^\perp$ and $\mathcal{N}^\perp$ are equal. More explicitly we can take the following steps:
- Let $v_1, v_2, \ldots, v_k$ be an ortho-normal basis of $\mathcal{M}$. Then $U_0 v_1, U_0 v_2, \ldots, U_0 v_k$ is an orthonormal basis of $\mathcal{N}$.
- Extend $\{v_1, v_2, \ldots, v_k\}$ to an orthonormal basis $\{v_1, v_2, \ldots, v_n\}$ of $\mathbb{C}^n$.

Continuation

- We can extend U_0 to a unitary from $\mathbb{C}^n$ to $\mathbb{C}^n$ as the dimensions of $\mathcal{M}^\perp$ and $\mathcal{N}^\perp$ are equal. More explicitly we can take the following steps:
- Let $v_1, v_2, \dots, v_k$ be an ortho-normal basis of $\mathcal{M}$. Then $U_0 v_1, U_0 v_2, \dots, U_0 v_k$ is an orthonormal basis of $\mathcal{N}$.
- Extend $\{v_1, v_2, \dots, v_k\}$ to an orthonormal basis $\{v_1, v_2, \dots, v_n\}$ of $\mathbb{C}^n$.
- Similarly, extend $\{U_0 v_1, U_0 v_2, \dots, U_0 v_k\}$ to an orthonormal basis $\{U_0 v_1, U_0 v_2, \dots, U_0 v_k, w_1, w_2, \dots, w_{n-k}\}$ of $\mathbb{C}^n$.

Continuation

- Now we extend U to a unitary of $\mathbb{C}^n$ by defining it on the orthonormal basis $\{v_1, \ldots, v_n\}$ (and extending linearly) by setting
$$Uv_j = \begin{cases} U_0 v_j & \text{for } 1 \leq j \leq k; \\ w_{j-k} & \text{for } k+1 \leq j \leq n. \end{cases}.$$

Continuation

- Now we extend U to a unitary of $\mathbb{C}^n$ by defining it on the orthonormal basis $\{v_1, \ldots, v_n\}$ (and extending linearly) by setting
$$Uv_j = \begin{cases} U_0 v_j & \text{for } 1 \leq j \leq k; \\ w_{j-k} & \text{for } k+1 \leq j \leq n. \end{cases}.$$
- Clearly U is a unitary and $Uv = U_0 v$ for $v \in \mathcal{M}$.

Continuation

- Now we extend U to a unitary of $\mathbb{C}^n$ by defining it on the orthonormal basis $\{v_1, \ldots, v_n\}$ (and extending linearly) by setting
$$Uv_j = \begin{cases} U_0 v_j & \text{for } 1 \leq j \leq k; \\ w_{j-k} & \text{for } k+1 \leq j \leq n. \end{cases}.$$
- Clearly U is a unitary and $Uv = U_0 v$ for $v \in \mathcal{M}$.
- Now for $x \in \mathbb{C}^n$,
$$U|B|x = URx = U_0 Rx = Bx, \quad \forall x \in \mathbb{C}^n.$$

Continuation

- Now we extend U to a unitary of $\mathbb{C}^n$ by defining it on the orthonormal basis $\{v_1, \ldots, v_n\}$ (and extending linearly) by setting
$$Uv_j = \begin{cases} U_0 v_j & \text{for } 1 \leq j \leq k; \\ w_{j-k} & \text{for } k+1 \leq j \leq n. \end{cases}.$$
- Clearly U is a unitary and $Uv = U_0 v$ for $v \in \mathcal{M}$.
- Now for $x \in \mathbb{C}^n$,
$$U|B|x = URx = U_0 Rx = Bx, \quad \forall x \in \mathbb{C}^n.$$
- This proves the existence. Now if $B = UR$ for some unitary U and positive R, $B^*B = R^*U^*UR = R^2$. Therefore R is the unique positive square root of B^*B, that is, $R = |B|$.

Continuation

- Now we extend U to a unitary of $\mathbb{C}^n$ by defining it on the orthonormal basis $\{v_1, \ldots, v_n\}$ (and extending linearly) by setting
$$Uv_j = \begin{cases} U_0 v_j & \text{for } 1 \leq j \leq k; \\ w_{j-k} & \text{for } k+1 \leq j \leq n. \end{cases}.$$
- Clearly U is a unitary and $Uv = U_0 v$ for $v \in \mathcal{M}$.
- Now for $x \in \mathbb{C}^n$,
$$U|B|x = URx = U_0 Rx = Bx, \quad \forall x \in \mathbb{C}^n.$$
- This proves the existence. Now if $B = UR$ for some unitary U and positive R, $B^*B = R^*U^*UR = R^2$. Therefore R is the unique positive square root of B^*B, that is, $R = |B|$.
- If B is invertible, and $B = UR$, then as $\det(B) = \det(U).\det(R) \neq 0$, R is also invertible and we get
$$U = BR^{-1} = B(|B|)^{-1}.$$

Continuation

- Now we extend U to a unitary of $\mathbb{C}^n$ by defining it on the orthonormal basis $\{v_1, \ldots, v_n\}$ (and extending linearly) by setting
$$Uv_j = \begin{cases} U_0 v_j & \text{for } 1 \leq j \leq k; \\ w_{j-k} & \text{for } k+1 \leq j \leq n. \end{cases}.$$
- Clearly U is a unitary and $Uv = U_0 v$ for $v \in \mathcal{M}$.
- Now for $x \in \mathbb{C}^n$,
$$U|B|x = URx = U_0 Rx = Bx, \quad \forall x \in \mathbb{C}^n.$$
- This proves the existence. Now if $B = UR$ for some unitary U and positive R, $B^*B = R^*U^*UR = R^2$. Therefore R is the unique positive square root of B^*B, that is, $R = |B|$.
- If B is invertible, and $B = UR$, then as $\det(B) = \det(U).\det(R) \neq 0$, R is also invertible and we get
$$U = BR^{-1} = B(|B|)^{-1}.$$
- Therefore, U is uniquely determined. ∎

Polar decomposition for normal matrices

- Example 28.2: Suppose A is a normal matrix and let

$$A = UDU^*$$

be the diagonalization of A with unitary U and diagonal matrix D. Let $d_1, d_2, \ldots, d_n$ be the diagonal entries of D.

Polar decomposition for normal matrices

- Example 28.2: Suppose A is a normal matrix and let
$$A = UDU^*$$
be the diagonalization of A with unitary U and diagonal matrix D. Let $d_1, d_2, \ldots, d_n$ be the diagonal entries of D.
- We can write D as
$$\begin{bmatrix} e^{i\theta_1} & 0 & \ldots & 0 \\ 0 & e^{i\theta_2} & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & e^{i\theta_n} \end{bmatrix} \cdot \begin{bmatrix} |d_1| & 0 & \ldots & 0 \\ 0 & |d_2| & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & |d_n| \end{bmatrix}.$$
using the polar decompositions of $d_1, d_2, \ldots, d_n$.

Polar decomposition for normal matrices

- Example 28.2: Suppose A is a normal matrix and let

$$A = UDU^*$$

be the diagonalization of A with unitary U and diagonal matrix D. Let $d_1, d_2, \ldots, d_n$ be the diagonal entries of D.

- We can write D as

$$\begin{bmatrix} e^{i\theta_1} & 0 & \ldots & 0 \\ 0 & e^{i\theta_2} & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & e^{i\theta_n} \end{bmatrix} \cdot \begin{bmatrix} |d_1| & 0 & \ldots & 0 \\ 0 & |d_2| & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & |d_n| \end{bmatrix}.$$

using the polar decompositions of $d_1, d_2, \ldots, d_n$.

- Note that if $d_j = 0$, then when we write $d_j = e^{i\theta_j}|d_j|$, $e^{i\theta_j}$ is not unique.

Continuation

- Now $A = V|A|$ where

$$V = U \begin{bmatrix} e^{i\theta_1} & 0 & \dots & 0 \\ 0 & e^{i\theta_2} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & e^{i\theta_n} \end{bmatrix} U^*$$

  is a unitary and

Continuation

- Now $A = V|A|$ where

$$V = U \begin{bmatrix} e^{i\theta_1} & 0 & \dots & 0 \\ 0 & e^{i\theta_2} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & e^{i\theta_n} \end{bmatrix} U^*$$

  is a unitary and

- $$|A| = U \begin{bmatrix} |d_1| & 0 & \dots & 0 \\ 0 & |d_2| & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & |d_n| \end{bmatrix} U^*$$

  is positive.

Continuation

- Alternatively, if $A = a_1 P_1 + a_2 P_2 + \cdots + a_k P_k$ is the spectral decomposition of A, we can get the polar decomposition as follows.

Continuation

- Alternatively, if $A = a_1P_1 + a_2P_2 + \cdots + a_kP_k$ is the spectral decomposition of A, we can get the polar decomposition as follows.
- Suppose $a_j = e^{i\theta_j}|a_j|, \quad 1 \leq j \leq k$ is the polar decomposition of a_j.

Continuation

- Alternatively, if $A = a_1 P_1 + a_2 P_2 + \cdots + a_k P_k$ is the spectral decomposition of A, we can get the polar decomposition as follows.
- Suppose $a_j = e^{i\theta_j}|a_j|, \quad 1 \leq j \leq k$ is the polar decomposition of a_j.
- Take $V = e^{i\theta_1} P_1 + e^{i\theta_2} P_2 + \cdots + e^{i\theta_k} P_k$ and $|A| = |a_1| P_1 + |a_2| P_2 + \cdots + |a_k| P_k$.

Continuation

- Alternatively, if $A = a_1 P_1 + a_2 P_2 + \cdots + a_k P_k$ is the spectral decomposition of A, we can get the polar decomposition as follows.
- Suppose $a_j = e^{i\theta_j}|a_j|, \quad 1 \leq j \leq k$ is the polar decomposition of a_j.
- Take $V = e^{i\theta_1} P_1 + e^{i\theta_2} P_2 + \cdots + e^{i\theta_k} P_k$ and $|A| = |a_1| P_1 + |a_2| P_2 + \cdots + |a_k| P_k$.
- Then $A = V|A|$ is the polar decomposition of A.

Continuation

- Alternatively, if $A = a_1 P_1 + a_2 P_2 + \cdots + a_k P_k$ is the spectral decomposition of A, we can get the polar decomposition as follows.
- Suppose $a_j = e^{i\theta_j}|a_j|, \quad 1 \leq j \leq k$ is the polar decomposition of a_j.
- Take $V = e^{i\theta_1}P_1 + e^{i\theta_2}P_2 + \cdots + e^{i\theta_k}P_k$ and $|A| = |a_1|P_1 + |a_2|P_2 + \cdots + |a_k|P_k$.
- Then $A = V|A|$ is the polar decomposition of A.
- END OF REVIEW

Singular value decomposition and singular values

- Theorem 29.1: Let B be an $n \times n$ complex matrix. Then there exist two unitaries V, W and a diagonal matrix S with non-negative entries such that

$$B = VSW.$$

Singular value decomposition and singular values

- Theorem 29.1: Let B be an $n \times n$ complex matrix. Then there exist two unitaries V, W and a diagonal matrix S with non-negative entries such that

$$B = VSW.$$

- Proof: By polar decomposition theorem,

$$B = U|B|$$

for some unitary U with $|B| = (B^*B)^{\frac{1}{2}}$.

Singular value decomposition and singular values

- Theorem 29.1: Let B be an $n \times n$ complex matrix. Then there exist two unitaries V, W and a diagonal matrix S with non-negative entries such that

$$B = VSW.$$

- Proof: By polar decomposition theorem,

$$B = U|B|$$

for some unitary U with $|B| = (B^*B)^{\frac{1}{2}}$.

- Now as $|B|$ is a normal operator there exists a unitary operator W and a diagonal matrix S such that

$$|B| = W^*SW.$$

Singular value decomposition and singular values

- Theorem 29.1: Let B be an $n \times n$ complex matrix. Then there exist two unitaries V, W and a diagonal matrix S with non-negative entries such that

$$B = VSW.$$

- Proof: By polar decomposition theorem,

$$B = U|B|$$

for some unitary U with $|B| = (B^*B)^{\frac{1}{2}}$.

- Now as $|B|$ is a normal operator there exists a unitary operator W and a diagonal matrix S such that

$$|B| = W^*SW.$$

- Taking, $V = UW^*$, V is a unitary and

$$B = VSW.$$

Singular value decomposition and singular values

- Theorem 29.1: Let B be an $n \times n$ complex matrix. Then there exist two unitaries V, W and a diagonal matrix S with non-negative entries such that
$$B = VSW.$$
- Proof: By polar decomposition theorem,
$$B = U|B|$$
for some unitary U with $|B| = (B^*B)^{\frac{1}{2}}$.
- Now as $|B|$ is a normal operator there exists a unitary operator W and a diagonal matrix S such that
$$|B| = W^*SW.$$
- Taking, $V = UW^*$, V is a unitary and
$$B = VSW.$$
- Note that as $|B|$ is positive, the diagonal entries of S are non-negative. ■

Continuation

- Definition 29.2: The decomposition of a matrix B as

$$B = VSW$$

where V, W are unitaries and S is diagonal with non-negative entries, is known as singular value decomposition. The diagonal entries of S are known as singular values of B.

Continuation

- Definition 29.2: The decomposition of a matrix B as

$$B = VSW$$

where V, W are unitaries and S is diagonal with non-negative entries, is known as singular value decomposition. The diagonal entries of S are known as singular values of B.

- The unitaries V, W in the singular value decomposition are not unique. However, as

$$B^*B = W^*S^*SW = W^*S^2W$$

the diagonal entries of S are eigenvalues of $|B|$. In other words, singular values of B are just eigenvalues of $|B|$.

Continuation

- Definition 29.2: The decomposition of a matrix B as

$$B = VSW$$

where V, W are unitaries and S is diagonal with non-negative entries, is known as singular value decomposition. The diagonal entries of S are known as singular values of B.

- The unitaries V, W in the singular value decomposition are not unique. However, as

$$B^*B = W^*S^*SW = W^*S^2W$$

the diagonal entries of S are eigenvalues of $|B|$. In other words, singular values of B are just eigenvalues of $|B|$.

- Note that,

$$BB^* = VSS^*V^* = VS^2V^*.$$

Continuation

- Definition 29.2: The decomposition of a matrix B as

$$B = VSW$$

where V, W are unitaries and S is diagonal with non-negative entries, is known as singular value decomposition. The diagonal entries of S are known as singular values of B.

- The unitaries V, W in the singular value decomposition are not unique. However, as

$$B^*B = W^*S^*SW = W^*S^2W$$

the diagonal entries of S are eigenvalues of $|B|$. In other words, singular values of B are just eigenvalues of $|B|$.

- Note that,

$$BB^* = VSS^*V^* = VS^2V^*.$$

- Therefore singular values of B are also eigenvalues of $(BB^*)^{\frac{1}{2}}$.

Continuation

- Further,
$$BB^* = VS^2V^* = VW(B^*B)W^*V^*.$$

Continuation

- Further,
$$BB^* = VS^2V^* = VW(B^*B)W^*V^*.$$
- Hence B^*B and BB^* are unitarily equivalent.

Eigenvector type equations

- Suppose $B = VSW$ is a singular value decomposition of B.

Eigenvector type equations

- Suppose $B = VSW$ is a singular value decomposition of B.
- Then, $B^* = W^*SV^*$ and so

$$BW^* = VS, \quad B^*V = W^*S.$$

Eigenvector type equations

- Suppose $B = VSW$ is a singular value decomposition of B.
- Then, $B^* = W^*SV^*$ and so

$$BW^* = VS, \quad B^*V = W^*S.$$

- Let $w_1, w_2, \ldots, w_n$ be the columns of W^* and $v_1, v_2, \ldots, v_n$ be the columns of V. Let $s_1, s_2, \ldots, s_n$ be the diagonal entries of S.

Eigenvector type equations

- Suppose $B = VSW$ is a singular value decomposition of B.
- Then, $B^* = W^*SV^*$ and so

$$BW^* = VS, \quad B^*V = W^*S.$$

- Let $w_1, w_2, \ldots, w_n$ be the columns of W^* and $v_1, v_2, \ldots, v_n$ be the columns of V. Let $s_1, s_2, \ldots, s_n$ be the diagonal entries of S.
- Then the matrix equations above can be written as:

$$Bw_j = s_j v_j, B^* v_j = s_j w_j, \quad 1 \leq j \leq n.$$

Eigenvector type equations

- Suppose $B = VSW$ is a singular value decomposition of B.
- Then, $B^* = W^*SV^*$ and so

$$BW^* = VS, \quad B^*V = W^*S.$$

- Let $w_1, w_2, \ldots, w_n$ be the columns of W^* and $v_1, v_2, \ldots, v_n$ be the columns of V. Let $s_1, s_2, \ldots, s_n$ be the diagonal entries of S.
- Then the matrix equations above can be written as:

$$Bw_j = s_j v_j, B^* v_j = s_j w_j, \quad 1 \leq j \leq n.$$

- So we have pairs of vectors, which behave somewhat like eigenvectors.

Eigenvector type equations

- Suppose $B = VSW$ is a singular value decomposition of B.
- Then, $B^* = W^*SV^*$ and so

$$BW^* = VS, \quad B^*V = W^*S.$$

- Let $w_1, w_2, \ldots, w_n$ be the columns of W^* and $v_1, v_2, \ldots, v_n$ be the columns of V. Let $s_1, s_2, \ldots, s_n$ be the diagonal entries of S.
- Then the matrix equations above can be written as:

$$Bw_j = s_j v_j, B^* v_j = s_j w_j, \quad 1 \leq j \leq n.$$

- So we have pairs of vectors, which behave somewhat like eigenvectors.
- It is possible to have singular value decomposition for non-square matrices. But we will not be considering that.

Eigenvector type equations

- Suppose $B = VSW$ is a singular value decomposition of B.
- Then, $B^* = W^*SV^*$ and so

$$BW^* = VS, \quad B^*V = W^*S.$$

- Let $w_1, w_2, \ldots, w_n$ be the columns of W^* and $v_1, v_2, \ldots, v_n$ be the columns of V. Let $s_1, s_2, \ldots, s_n$ be the diagonal entries of S.
- Then the matrix equations above can be written as:

$$Bw_j = s_j v_j, B^* v_j = s_j w_j, \quad 1 \leq j \leq n.$$

- So we have pairs of vectors, which behave somewhat like eigenvectors.
- It is possible to have singular value decomposition for non-square matrices. But we will not be considering that.
- The notion of SVD has many practical applications.

Normal matrices

- If A is normal, then by diagonalization,

$$A = UDU^*$$

for some unitary U and diagonal matrix D.

Normal matrices

- If A is normal, then by diagonalization,

$$A = UDU^*$$

  for some unitary U and diagonal matrix D.
- Let $d_j = e^{i\theta_j}|d_j|$ be the diagonal entries of D, $1 \leq j \leq n$.

Normal matrices

- If A is normal, then by diagonalization,

$$A = UDU^*$$

  for some unitary U and diagonal matrix D.
- Let $d_j = e^{i\theta_j}|d_j|$ be the diagonal entries of D, $1 \leq j \leq n$.
- Then the singular value decomposition of A is given by

$$A = V_1 \begin{bmatrix} |d_1| & 0 & \dots & 0 \\ 0 & |d_2| & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & |d_n| \end{bmatrix} U^*,$$

Normal matrices

- If A is normal, then by diagonalization,

$$A = UDU^*$$

  for some unitary U and diagonal matrix D.
- Let $d_j = e^{i\theta_j}|d_j|$ be the diagonal entries of D, $1 \le j \le n$.
- Then the singular value decomposition of A is given by

$$A = V_1 \begin{bmatrix} |d_1| & 0 & \dots & 0 \\ 0 & |d_2| & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & |d_n| \end{bmatrix} U^*,$$

- where

$$V_1 = U \begin{bmatrix} e^{i\theta_1} & 0 & \dots & 0 \\ 0 & e^{i\theta_2} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & e^{i\theta_n} \end{bmatrix}.$$

Example

- We revisit an example we saw before.

Example

- We revisit an example we saw before.
- Example 28.3 and 29.4: Consider

$$B = \begin{bmatrix} 2 & 2 & 2 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

- Then

$$B^* B = \begin{bmatrix} 4 & 4 & 4 \\ 4 & 4 & 4 \\ 4 & 4 & 4 \end{bmatrix}.$$

Example

- We revisit an example we saw before.
- Example 28.3 and 29.4: Consider

$$B = \begin{bmatrix} 2 & 2 & 2 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

- Then

$$B^*B = \begin{bmatrix} 4 & 4 & 4 \\ 4 & 4 & 4 \\ 4 & 4 & 4 \end{bmatrix}.$$

- Recall that an $n \times n$ matrix with all entries equal to $\frac{1}{n}$ is a projection.

Example

- We revisit an example we saw before.
- Example 28.3 and 29.4: Consider

$$B = \begin{bmatrix} 2 & 2 & 2 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

- Then

$$B^*B = \begin{bmatrix} 4 & 4 & 4 \\ 4 & 4 & 4 \\ 4 & 4 & 4 \end{bmatrix}.$$

- Recall that an $n \times n$ matrix with all entries equal to $\frac{1}{n}$ is a projection.
- Therefore

$$B^*B = 12 \begin{bmatrix} 1/3 & 1/3 & 1/3 \\ 1/3 & 1/3 & 1/3 \\ 1/3 & 1/3 & 1/3 \end{bmatrix} + 0. \begin{bmatrix} 2/3 & -1/3 & -1/3 \\ -1/3 & 2/3 & -1/3 \\ -1/3 & -1/3 & 2/3 \end{bmatrix}$$

is the spectral decomposition of B^*B.

Continuation

- Therefore,

$$\begin{aligned} |B| &= \sqrt{12}\begin{bmatrix} 1/3 & 1/3 & 1/3 \\ 1/3 & 1/3 & 1/3 \\ 1/3 & 1/3 & 1/3 \end{bmatrix} \\ &= \begin{bmatrix} 2/\sqrt{3} & 2/\sqrt{3} & 2/\sqrt{3} \\ 2/\sqrt{3} & 2/\sqrt{3} & 2/\sqrt{3} \\ 2/\sqrt{3} & 2/\sqrt{3} & 2/\sqrt{3} \end{bmatrix}. \end{aligned}$$

Continuation

- Take
$$U = \begin{bmatrix} 1/\sqrt{3} & 1/\sqrt{3} & 1/\sqrt{3} \\ 1/\sqrt{2} & -1/\sqrt{2} & 0 \\ 1/\sqrt{6} & 1/\sqrt{6} & -2/\sqrt{6} \end{bmatrix}.$$

Continuation

- Take
$$U = \begin{bmatrix} 1/\sqrt{3} & 1/\sqrt{3} & 1/\sqrt{3} \\ 1/\sqrt{2} & -1/\sqrt{2} & 0 \\ 1/\sqrt{6} & 1/\sqrt{6} & -2/\sqrt{6} \end{bmatrix}.$$
- We may now verify that
$$B = \begin{bmatrix} 1/\sqrt{3} & 1/\sqrt{3} & 1/\sqrt{3} \\ 1/\sqrt{2} & -1/\sqrt{2} & 0 \\ 1/\sqrt{6} & 1/\sqrt{6} & -2/\sqrt{6} \end{bmatrix} \cdot \begin{bmatrix} 2/\sqrt{3} & 2/\sqrt{3} & 2/\sqrt{3} \\ 2/\sqrt{3} & 2/\sqrt{3} & 2/\sqrt{3} \\ 2/\sqrt{3} & 2/\sqrt{3} & 2/\sqrt{3} \end{bmatrix}.$$

  is a polar decomposition of B.

Continuation

- $$B = I.\begin{bmatrix} 2\sqrt{3} & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}.\begin{bmatrix} 1/\sqrt{3} & 1/\sqrt{3} & 1/\sqrt{3} \\ 1/\sqrt{2} & -1/\sqrt{2} & 0 \\ 1/\sqrt{6} & 1/\sqrt{6} & -2/\sqrt{6} \end{bmatrix},$$

  is a singular value decomposition of B.

Continuation

- $$B = I . \begin{bmatrix} 2\sqrt{3} & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} . \begin{bmatrix} 1/\sqrt{3} & 1/\sqrt{3} & 1/\sqrt{3} \\ 1/\sqrt{2} & -1/\sqrt{2} & 0 \\ 1/\sqrt{6} & 1/\sqrt{6} & -2/\sqrt{6} \end{bmatrix},$$

  is a singular value decomposition of B.
-

Continuation

- $$B = I.\begin{bmatrix} 2\sqrt{3} & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}.\begin{bmatrix} 1/\sqrt{3} & 1/\sqrt{3} & 1/\sqrt{3} \\ 1/\sqrt{2} & -1/\sqrt{2} & 0 \\ 1/\sqrt{6} & 1/\sqrt{6} & -2/\sqrt{6} \end{bmatrix},$$

  is a singular value decomposition of B.
-
- END OF LECTURE 29.