CS724: Topics in Algorithms
Matrices
Slide Set 3

Prof. Dan A. Simovici

1. Matrices with Real and Complex Elements
2. Complex Matrices
3. Partitioned Matrices and Matrix Operations
4. Invertible matrices
5. Change of Bases
6. Matrices and Bilinear Forms
7. Matrices and Linear Transformations
8. The Notion of Rank
9. Matrix Similarity and Congruence
10. The Kronecker and Hadamard products
11. Outer Products

The $(n \times n)$-*unit matrix* on the field $\mathbb{F}$ is the square matrix $I_n \in \mathbb{F}^{n\times n}$ given by

$$I_n = \begin{pmatrix} 1 & 0 & 0 & \cdots & 0 \\ 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \end{pmatrix}.$$

The $(m \times n)$-*zero matrix* is the $(m \times n)$-matrix $O_{m,n} \in \mathbb{F}^{n\times n}$ given by

$$O_{m,n} = \begin{pmatrix} 0 & 0 & 0 & \cdots & 0 \\ 0 & 0 & 0 & \cdots & 0 \\ 0 & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & 0 & \cdots & 0 \end{pmatrix}.$$

The $(m \times n)$-*complete matrix* is the $(m \times n)$-matrix $J_{m,n}$ given by

$$J_{m,n} = \begin{pmatrix} 1 & 1 & 1 & \cdots & 1 \\ 1 & 1 & 1 & \cdots & 1 \\ 1 & 1 & 1 & \cdots & 1 \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ 1 & 1 & 1 & \cdots & 1 \end{pmatrix}.$$

The one-column matrix $O_{m,1}$ is denoted by $\mathbf{0}_m$. Similarly, the one column matrix having m rows

$$\begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix}$$

is denoted by $\mathbf{1}_m$. The subscripts are omitted whenever there is no ambiguity.

Definition

A *diagonal matrix* is a matrix $D \in \mathbb{F}^{m\times n}$ such that $i \neq j$ implies $d_{ij} = 0$. If $p \leqslant \min\{m, n\}$, then we denote the diagonal matrix

$$D = \begin{pmatrix} d_1 & 0 & 0 & 0 & \cdots & 0 \\ 0 & d_2 & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \ldots & \vdots \\ 0 & 0 & 0 & 0 & \cdots & 0 \\ 0 & 0 & 0 & 0 & \cdots & 0 \end{pmatrix}$$

by $\mathrm{diag}(d_1, \ldots, d_p)$.

Definition

Let $A = (a_{ij})$ and $B = (b_{ij})$ be two matrices in $\mathbb{F}^{m \times n}$.
The *sum* of the matrices A and B is the matrix $A + B$ having the same format and defined by

$$(A + B)(i, j) = a_{ij} + b_{ij}$$

for $1 \leqslant i \leqslant m$ and $1 \leqslant j \leqslant n$.

It is easy to verify that the matrix sum is an associative and commutative operation on $\mathbb{F}^{m\times n}$; that is,

$$\begin{aligned} A+(B+C) &= (A+B)+C, \\ A+B &= B+A, \end{aligned}$$

for all $A, B, C \in \mathbb{F}^{m\times n}$.
The zero matrix $O_{m,n}$ acts as an additive unit on the set $S^{m\times n}$; that is,

$$A + O_{m,n} = O_{m,n} + A,$$

for every $A \in \mathbb{F}^{m\times n}$.
The additive inverse, or the *opposite of a matrix* $A = (a_{ij}) \in \mathbb{F}^{m\times n}$, is the matrix $-A$ given by $(-A)(i,j) = -a_{ij}$ for $1 \leqslant i \leqslant m$ and $1 \leqslant j \leqslant n$.

Example

The opposite of $A \in \mathbb{R}^{2\times 3}$, given by

$$A = \begin{pmatrix} 1 & -2 & 3 \\ 0 & 2 & -1 \end{pmatrix}$$

is the matrix

$$-A = \begin{pmatrix} -1 & 2 & -3 \\ 0 & -2 & 1 \end{pmatrix}.$$

It is immediate that $A + (-A) = O_{2,3}$.

UMASS BOSTON

Definition

Let $A \in \mathbb{F}^{m\times n}$ and $B \in \mathbb{F}^{n\times p}$ be two matrices. The *product* of the matrices A, B is the matrix $C \in \mathbb{F}^{m\times p}$ defined by

$$C(i,k) = \sum_{j=1}^{n} a_{ij} b_{jk},$$

where $1 \leqslant i \leqslant m$ and $1 \leqslant k \leqslant p$. The product of the matrices A, B is denoted by AB.

Matrix multiplication of A and B is possible only if number of columns of A is equal to the number of rows of the second matrix B. Any pair of matrices (A, B) that satisfies this condition is said to be *conformant*.

Theorem

Matrix multiplication is associative, that is, $A(BC) = (AB)C$.

Theorem

If $A \in \mathbb{F}^{m \times n}$, *then* $I_m A = A I_n = A$.

The product of matrices is not commutative. Indeed, consider the matrices $A, B \in \mathbb{Z}^{2\times 2}$ defined by

$$A = \begin{pmatrix} 0 & 1 \\ 2 & 3 \end{pmatrix} \text{ and } B = \begin{pmatrix} -1 & 1 \\ 1 & 0 \end{pmatrix}.$$

We have

$$AB = \begin{pmatrix} 1 & 0 \\ 1 & 2 \end{pmatrix} \text{ and } BA = \begin{pmatrix} 2 & 2 \\ 0 & 1 \end{pmatrix},$$

so $AB \neq BA$.

Definition

A matrix $A \in \mathbb{F}^{n \times n}$ is *upper triangular* if $j < i$ implies $a_{ij} = 0$ and is *strictly upper triangular* if $j \leqslant i$ implies $a_{ij} = 0$.
A is *lower triangular* (*strictly lower triangular*) if A' is upper triangular (strictly upper triangular).

Example

The matrix $L \in \mathbb{Z}^{4\times 4}$ given by

$$L = \begin{pmatrix} 1 & 0 & 0 & 0 \\ -1 & 2 & 0 & 0 \\ 3 & 1 & 2 & 0 \\ -7 & 6 & 1 & -6 \end{pmatrix}$$

is a lower triangular matrix. The matrix $U \in \mathbb{Z}^{4\times 4}$

$$U = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 0 & -1 & 2 & 3 \\ 0 & 0 & -1 & -2 \\ 0 & 0 & 0 & 2 \end{pmatrix}$$

is an upper triangular matrix.

It is interesting to compute two matrix products that can be formed starting from the columns $\boldsymbol{u}$ and $\boldsymbol{v}$ given by

$$\boldsymbol{u} = \begin{pmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{pmatrix} \text{ and } \boldsymbol{v} = \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix}.$$

Note that $\boldsymbol{u}'\boldsymbol{v} \in \mathbb{F}^{1\times 1}$, that is,

$$\boldsymbol{u}'\boldsymbol{v} = u_1 v_1 + u_2 v_2 + \cdots + u_n v_n = \boldsymbol{v}'\boldsymbol{u}.$$

This product is known as the *inner product* of $\boldsymbol{u}$ and $\boldsymbol{v}$.

Let $\mathbb{C}$ be the field of complex numbers. A *complex* matrix is a matrix $A \in \mathbb{C}^{m \times n}$.

Definition

The conjugate of a matrix $A \in \mathbb{C}^{m \times n}$ is the matrix $\bar{A} \in \mathbb{C}^{m \times n}$, where $\bar{A}(i,j) = \overline{A(i,j)}$ for $1 \leqslant i \leqslant m$ and $1 \leqslant j \leqslant n$.

The notion of symmetry is extended to accommodate complex matrices.

Definition

The *transpose conjugate* of the matrix $A \in \mathbb{C}^{m \times n}$ or its *Hermitian adjoint* is the matrix $B \in \mathbb{C}^{n \times m}$ given by $B = \bar{A}' = \overline{(A')}$.

The transpose conjugate of A is denoted by A^{H}.

Example

Let $A \in \mathbb{C}^{3\times 2}$ be the matrix

$$A = \begin{pmatrix} 1+i & 2 \\ 2-i & i \\ 0 & 1-2i \end{pmatrix}.$$

The matrix A^{H} is given by

$$A^{\mathsf{H}} = \begin{pmatrix} 1-i & 2+i & 0 \\ 2 & -i & 1+2i \end{pmatrix}.$$

Using Hermitian conjugates several important classes of matrices are defined.

Definition

The matrix $A \in \mathbb{C}^{n \times n}$ is:

- *Hermitian* if $A = A^{\mathsf{H}}$;
- *skew-Hermitian* if $A^{\mathsf{H}} = -A$;
- *normal* if $AA^{\mathsf{H}} = A^{\mathsf{H}}A$;
- *unitary* if $AA^{\mathsf{H}} = A^{\mathsf{H}}A = I_n$.

Example

Let $\alpha, \beta, \gamma, \delta$ and θ be five real numbers such that $\alpha - \beta - \gamma + \delta$ is a multiple of 2π. The matrix

$$M_{\alpha,\beta,\gamma,\delta}(\theta) = \begin{pmatrix} e^{i\alpha}\cos\theta & -e^{i\beta}\sin\theta \\ e^{i\gamma}\sin\theta & e^{i\delta}\cos\theta \end{pmatrix}$$

is unitary because

$$\begin{aligned} & M_{\alpha,\beta,\gamma,\delta}(\theta)^{\mathsf{H}} M_{\alpha,\beta,\gamma,\delta}(\theta) \\ = & \begin{pmatrix} e^{-i\alpha}\cos\theta & e^{-i\gamma}\sin\theta \\ -e^{-i\beta}\sin\theta & e^{-i\delta}\cos\theta \end{pmatrix} \begin{pmatrix} e^{i\alpha}\cos\theta & -e^{i\beta}\sin\theta \\ e^{i\gamma}\sin\theta & e^{i\delta}\cos\theta \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}. \end{aligned}$$

The following properties of unitary matrices can be easily verified:

- all unitary matrices are normal;
- a matrix $A \in \mathbb{C}^{n \times n}$ is unitary if and only if A^{H} is unitary;
- the product of two unitary matrices is an unitary matrix.

We verify only the last property. Suppose that $A, B \in \mathbb{C}^{n\times n}$ are unitary matrices, that is, $AA^{\mathsf{H}} = BB^{\mathsf{H}} = I_n$. Then
$(AB)(AB)^{\mathsf{H}} = ABB^{\mathsf{H}}A^{\mathsf{H}} = AA^{\mathsf{H}} = I_n$, hence AB is unitary.
If $A \in \mathbb{R}^{n\times n}$ is a real matrix and A is unitary we refer to A as an *orthogonal matrix* or an *orthonormal matrix*.

If $A \in \mathbb{R}^{n \times n}$ is a matrix with real entries, then its Hermitian adjoint coincides with the transposed matrix A'. Thus, a real matrix is Hermitian if and only if it is symmetric.

Observe that if $\boldsymbol{z} \in \mathbb{C}^n$ and

$$\boldsymbol{z} = \begin{pmatrix} z_1 \\ \vdots \\ z_n \end{pmatrix},$$

then $\boldsymbol{z}^{\mathsf{H}}\boldsymbol{z} = \overline{z}_1 z_1 + \cdots + \overline{z}_n z_n = \sum_{i=1}^{n} |z_i|^2$.

Theorem

Let $A \in \mathbb{C}^{n\times n}$. The following statements hold:

- *the matrices $A + A^H$, AA^H and $A^H A$ are Hermitian and $A - A^H$ is skew-Hermitian;*
- *if A is a Hermitian matrix, then so is A^k for $k \in \mathbb{N}$;*
- *if A is Hermitian and invertible, then so is A^{-1};*
- *if A is Hermitian, then a_{ii} are real numbers for $1 \leqslant i \leq n$.*

Proof.

All statements follow directly from the definition of Hermitian matrices. □

Theorem

If $A \in \mathbb{C}^{n\times n}$ there exists a unique pair of Hermitian matrices (H_1, H_2) such that $A = H_1 + iH_2$.

Proof.

Let

$$H_1 = \frac{1}{2}(A + A^{\mathsf{H}}) \text{ and } H_2 = -\frac{i}{2}(A - A^{\mathsf{H}}).$$

It is immediate that both H_1 and H_2 are Hermitian and that $H_1 + iH_2 = A$. Suppose that $A = H_3 + iH_4$, where H_3 and H_4 are Hermitian. Then, we have

$$\begin{aligned} 2H_1 &= A + A^{\mathsf{H}} = H_3 + iH_4 + H_3^{\mathsf{H}} - iH_4^{\mathsf{H}} \\ &= 2H_3, \end{aligned}$$

so $H_1 = H_3$. Therefore $H_2 = H_4$, so the matrices H_1 and H_2 are uniquely determined. □

Theorem

If $A \in \mathbb{C}^{n\times n}$ there exists a unique pair of matrices (H, S) such that H is Hermitian, S is skew-Hermitian and $A = H + S$.

Proof.

A can be written as $A = H_1 + iH_2$, where H_1 and H_2 are Hermitian matrices. Choose $H = H_1$ and $S = iH_2$. S is skew-Hermitian. The uniqueness of the pair (H, S) is immediate. □

Next, we discuss a characterization of Hermitian matrices.

Theorem

A matrix $A \in \mathbb{C}^{n\times n}$ is Hermitian if and only if $\mathbf{x}^{\mathsf{H}} A\mathbf{x}$ is a real number for every $\mathbf{x} \in \mathbb{C}^n$.

Proof: Suppose that A is Hermitian. Then,

$$\overline{\mathbf{x}^{\mathsf{H}} A\mathbf{x}} = \overline{\mathbf{x}^{\mathsf{H}} A^{\mathsf{H}}\mathbf{x}} = \mathbf{x}' A' \overline{\mathbf{x}} = \mathbf{x}' A' (\mathbf{x}^{\mathsf{H}})' = \mathbf{x}^{\mathsf{H}} A\mathbf{x},$$

so $\mathbf{x}^{\mathsf{H}} A\mathbf{x}$ is a real number because it is equal to its conjugate.

Proof cont'd:

Conversely, suppose that $\mathbf{x}^{\mathsf{H}} A \mathbf{x}$ is a real number for every $\mathbf{x} \in \mathbb{C}^n$. This implies that

$$(\mathbf{x}+\mathbf{y})^{\mathsf{H}} A(\mathbf{x}+\mathbf{y}) = \mathbf{x}^{\mathsf{H}} A \mathbf{x} + \mathbf{x}^{\mathsf{H}} A \mathbf{y} + \mathbf{y}^{\mathsf{H}} A \mathbf{x} + \mathbf{y}^{\mathsf{H}} A \mathbf{y}$$

is a real number, so $\mathbf{x}^{\mathsf{H}} A \mathbf{y} + \mathbf{y}^{\mathsf{H}} A \mathbf{x}$ is real for every $\mathbf{x}, \mathbf{y} \in \mathbb{C}^n$.
Let $\mathbf{x} = \mathbf{e}_p$ and $\mathbf{y} = \mathbf{e}_q$. Then, $a_{pq} + a_{qp}$ is a real number. If we choose $\mathbf{x} = -i\mathbf{e}_p$ and $\mathbf{y} = \mathbf{e}_j$ it follows that $-ia_{pq} + ia_{qp}$ is a real number. Thus, $\Im(a_{pq}) = -\Im(a_{qp})$ and $\Re(a_{pq}) = \Re(a_{qp})$, which leads to $a_{pq} = \overline{a}_{qp}$ for $1 \leqslant p, q \leqslant n$. These equalities are equivalent to $A = A^{\mathsf{H}}$, so A is Hermitian.

Let $A \in \mathbb{F}^{m \times n}$ be a matrix and suppose that $m = m_1 + \cdots + m_p$ and $n = n_1 + \cdots + n_q$, where $\mathbb{F}$ is the real or the complex field.
A *partitioning of A* is a collection of matrices $A_{hk} \in \mathbb{F}^{m_h \times n_k}$ such that A_{hk} is the contiguous submatrix

$$A\left[\begin{array}{c} m_1 + \cdots + m_{h-1} + 1, \ldots, m_1 + \cdots + m_{h-1} + m_h \\ n_1 + \cdots + n_{k-1} + 1, \ldots, n_1 + \cdots + n_k \end{array}\right],$$

for $1 \leqslant h \leqslant p$ and $1 \leqslant k \leqslant q$.
If $\{A_{hk} \mid 1 \leqslant h \leqslant p \text{ and } 1 \leqslant k \leqslant q\}$ is a partitioning of A, A is written as

$$A = \begin{pmatrix} A_{11} & A_{12} & \cdots & A_{1q} \\ A_{21} & A_{22} & \cdots & A_{2q} \\ \vdots & \vdots & \cdots & \vdots \\ A_{p1} & A_{p2} & \cdots & A_{pq} \end{pmatrix}.$$

UMASS BOSTON

The matrices A_{hk} are referred to as the *blocks* of the partitioning. All blocks located in a column must have the number of columns; all blocks located in a row must have the same number of rows.
The matrix $A \in \mathbb{F}^{5\times 6}$ given by

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} & a_{14} & a_{15} & a_{16} \\ a_{21} & a_{22} & a_{23} & a_{24} & a_{25} & a_{26} \\ a_{31} & a_{32} & a_{33} & a_{34} & a_{35} & a_{36} \\ a_{41} & a_{42} & a_{43} & a_{44} & a_{45} & a_{46} \\ a_{51} & a_{52} & a_{53} & a_{54} & a_{55} & a_{56} \end{pmatrix}$$

can be partitioned as

$$\left(\begin{array}{ccc|c|cc} a_{11} & a_{12} & a_{13} & a_{14} & a_{15} & a_{16} \\ a_{21} & a_{22} & a_{23} & a_{24} & a_{25} & a_{26} \\ a_{31} & a_{32} & a_{33} & a_{34} & a_{35} & a_{36} \\ \hline a_{41} & a_{42} & a_{43} & a_{44} & a_{45} & a_{46} \\ a_{51} & a_{52} & a_{53} & a_{54} & a_{55} & a_{56} \end{array}\right)$$

Thus, if we introduce the matrices

$$A_{11} = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}, \quad A_{12} = \begin{pmatrix} a_{14} \\ a_{24} \\ a_{34} \end{pmatrix}, \quad A_{13} = \begin{pmatrix} a_{15} & a_{16} \\ a_{25} & a_{26} \\ a_{35} & a_{36} \end{pmatrix},$$

$$A_{21} = \begin{pmatrix} a_{41} & a_{42} & a_{43} \\ a_{51} & a_{52} & a_{53} \end{pmatrix}, \quad A_{22} = \begin{pmatrix} a_{45} \\ a_{55} \end{pmatrix}, \quad A_{23} = \begin{pmatrix} a_{45} & a_{46} \\ a_{55} & a_{56} \end{pmatrix},$$

the matrix A can be written as

$$A = \begin{pmatrix} A_{11} & A_{12} & A_{13} \\ A_{21} & A_{22} & A_{23} \end{pmatrix}.$$

Partitioning matrices is useful because matrix operations can be performed on block submatrices in a manner similar to scalar operations as we show next.

Theorem

Let $A \in \mathbb{F}^{m \times n}$ and $B \in \mathbb{F}^{n \times p}$ be two matrices. Suppose that the matrices A, B are partitioned as

$$A = \begin{pmatrix} A_{11} & \cdots & A_{1k} \\ \vdots & \ldots & \vdots \\ A_{h1} & \cdots & A_{hk} \end{pmatrix} \text{ and } B = \begin{pmatrix} B_{11} & \cdots & B_{1\ell} \\ \vdots & \ldots & \vdots \\ B_{k1} & \cdots & B_{k\ell} \end{pmatrix},$$

where $A_{rs} \in \mathbb{F}^{m_r \times n_s}$, $B_{st} \in \mathbb{F}^{n_s \times p_t}$ for $1 \leqslant r \leqslant h$, $1 \leqslant s \leqslant k$ and $1 \leqslant t \leqslant \ell$. Then, the product $C = AB$ can be partitioned as

$$C = \begin{pmatrix} C_{11} & \ldots & C_{1\ell} \\ \vdots & \ldots & \vdots \\ C_{h1} & \cdots & C_{hl} \end{pmatrix},$$

where $C_{uv} = \sum_{t=1}^{k} A_{ut} B_{tv}$, $1 \leqslant u \leqslant h$, and $1 \leqslant v \leqslant \ell$.

Definition

A matrix $A \in \mathbb{C}^{n \times n}$ is invertible if there exists a matrix $B \in \mathbb{C}^{n \times n}$ such that $AB = BA = I_n$.

Theorem

If $A, B \in \mathbb{C}^{n\times n}$ *are two invertible matrices, then the product* AB *is invertible and* $(AB)^{-1} = B^{-1}A^{-1}$.
If $A \in \mathbb{C}^{n\times n}$ *is invertible, then* A^H *is invertible and* $(A^H)^{-1} = (A^{-1})^H$.

Theorem

Let $\{\mathbf{r}_1, \ldots, \mathbf{r}_n\}$ *be a basis in* $\mathbb{C}^n$.
A matrix $A \in \mathbb{C}^{n \times n}$ *is invertible if and only if the set of vectors* $\{A\mathbf{r}_1, \ldots, A\mathbf{r}_n\}$ *is a basis in* $\mathbb{C}^n$.

Proof

Suppose that A is an invertible matrix. Note that $A\mathbf{x}_i = A\mathbf{x}_j$ implies $\mathbf{x}_i = \mathbf{x}_j$, so $\{A\mathbf{r}_1, \ldots, A\mathbf{r}_n\}$ consists on n distinct vectors. We claim that the set $\{A\mathbf{r}_1, \ldots, A\mathbf{r}_n\}$ is linearly independent. Indeed, suppose that $c_1 A\mathbf{r}_1 + \cdots + c_n A\mathbf{r}_n = \mathbf{0}_n$ such that not all coefficients c_i equal 0. Then, by multiplying by A^{-1} to the left we obtain $c_1\mathbf{r}_1 + \cdots + c_n\mathbf{r}_n = \mathbf{0}_n$, which contradicts the fact that $\{\mathbf{r}_1, \ldots, \mathbf{r}_n\}$ is a basis. Thus, $\{A\mathbf{r}_1, \ldots, A\mathbf{r}_n\}$ is a linearly independent that consists of n vectors, which means that this set is a basis in $\mathbb{C}^n$.

Proof cont'd

Conversely, suppose that for any basis $\{\boldsymbol{r}_1, \ldots, \boldsymbol{r}_n\}$ of $\mathbb{C}^n$ the set $\{A\boldsymbol{r}_1, \ldots, A\boldsymbol{r}_n\}$ is a basis in $\mathbb{C}^n$. Each of the vectors $\boldsymbol{r}_i$ can be uniquely expressed as a linear combination of $A\boldsymbol{r}_1, \ldots, A\boldsymbol{r}_n$. In particular, for the standard basis $\{\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n\}$, each of the vectors $\boldsymbol{e}_i$ can be uniquely expressed as a linear combination of the vectors $A\boldsymbol{e}_1 = \boldsymbol{a}_1, \ldots, A\boldsymbol{e}_n = \boldsymbol{a}_n$, where $\boldsymbol{a}_1, \ldots, \boldsymbol{a}_n$ are the columns of the matrix A. In other words, we have the equalities

$$\boldsymbol{e}_i = b_{i1}\boldsymbol{a}_1 + \cdots + b_{in}\boldsymbol{a}_n$$

for $1 \leqslant i \leqslant n$. In a succinct form, these equalities can be written as $I_n = BA$, where B is the matrix of the coefficients b_{ij}, which shows that A is an invertible matrix.

Theorem

Let $\mathbf{e}_1, \ldots, \mathbf{e}_n$ *and* $\tilde{\mathbf{e}}_1, \ldots, \tilde{\mathbf{e}}_n$ *be two bases of an* $\mathbb{F}$*-linear space* L*. There exists an invertible matrix* $P \in \mathbb{F}^{n\times n}$ *such that*

$$(\mathbf{e}_1 \ \cdots \ \mathbf{e}_n) = (\tilde{\mathbf{e}}_1 \ \cdots \ \tilde{\mathbf{e}}_n)P.$$

Since $\tilde{\boldsymbol{e}}_1, \ldots, \tilde{\boldsymbol{e}}_n$ is a basis of L each vector $\boldsymbol{e}_i$ is a unique linear combination of the vectors $\tilde{\boldsymbol{e}}_1, \ldots, \tilde{\boldsymbol{e}}_n$, that is

$$\boldsymbol{e}_i = p_{1i}\tilde{\boldsymbol{e}}_1 + \cdots + p_{ni}\tilde{\boldsymbol{e}}_n = (\tilde{\boldsymbol{e}}_1 \ \cdots \ \tilde{\boldsymbol{e}}_n) \begin{pmatrix} p_{1i} \\ \vdots \\ p_{ni} \end{pmatrix},$$

for $1 \leqslant i \leqslant n$, so the equality of the theorem holds for the matrix $P = (p_{ij})$. We have to show that P is an invertible matrix. Assume that $P\boldsymbol{t} = \mathbf{0}_L$. The equality of the theorem implies

$$(\boldsymbol{e}_1 \ \cdots \ \boldsymbol{e}_n) \begin{pmatrix} t_1 \\ \vdots \\ t_n \end{pmatrix} = (\tilde{\boldsymbol{e}}_1 \ \cdots \ \tilde{\boldsymbol{e}}_n) P\boldsymbol{t} = \mathbf{0}_L.$$

which implies $t_1\boldsymbol{e}_1 + \cdots + t_n\boldsymbol{e}_n = \mathbf{0}_L$. Since $\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n$ is a basis we obtain $t_1 = \cdots = t_n = 0$, so $\boldsymbol{t} = \mathbf{0}_L$, which implies that P is an invertible matrix.

Corollary

Let $\mathbf{z} \in L$ and assume that $\mathbf{z}$ can be expressed relatively to the bases $\{\mathbf{e}_1, \ldots, \mathbf{e}_n\}$ and $\{\tilde{\mathbf{e}}_1, \ldots, \tilde{\mathbf{e}}_n\}$ as

$$\mathbf{z} = \sum_{i=1}^{n} x_i \mathbf{e}_i = \sum_{i=1}^{n} y_i \tilde{\mathbf{e}}_i,$$

respectively. If $P \in \mathbb{F}^{n \times n}$ is a matrix such that $(\mathbf{e}_1 \ \cdots \ \mathbf{e}_n) = (\tilde{\mathbf{e}}_1 \ \cdots \ \tilde{\mathbf{e}}_n)P$, then

$$P \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}.$$

Proof

We have

$$\mathbf{z} = (\mathbf{e}_1 \ \cdots \ \mathbf{e}_n) \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = (\tilde{\mathbf{e}}_1 \ \cdots \ \tilde{\mathbf{e}}_n) P \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}.$$

Substituting $(\mathbf{e}_1 \ \cdots \ \mathbf{e}_n)$ in the previous equality yields:

$$(\tilde{\mathbf{e}}_1 \ \cdots \ \tilde{\mathbf{e}}_n) P \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}.$$

Since $\{\tilde{\mathbf{e}}_1, \ldots, \tilde{\mathbf{e}}_n\}$ is a basis we obtain the equality

$$P \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}.$$

Thus, the components of a vector $\boldsymbol{z}$ relative to the two bases $(\boldsymbol{e}_1 \ \cdots \ \boldsymbol{e}_n)$, and $(\tilde{\boldsymbol{e}}_1 \ \cdots \ \tilde{\boldsymbol{e}}_n)$ transform in opposite direction to the basis transformation. We say that the set of numbers $\{x_1, \ldots, x_n\}$ are *contravariant* components of the vector $\boldsymbol{z}$.

We proved that if $\{\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n\}$ is a basis of the $\mathbb{F}$-linear space L, then the set of linear forms $\{\boldsymbol{f}^j \mid 1 \leqslant j \leqslant n\}$ defined by

$$\boldsymbol{f}^j(\boldsymbol{e}_i) = \begin{cases} 1 & \text{if } i = j, \\ 0 & \text{otherwise} \end{cases}$$

is a basis of the dual space L^*. Furthermore, if $\boldsymbol{f}(\boldsymbol{e}_i) = a_i$, then $\boldsymbol{f} = a_1\boldsymbol{f}^1 + \cdots + a_n\boldsymbol{f}^n$.

If we change the basis in L such that $(\boldsymbol{e}_1 \ \cdots \ \boldsymbol{e}_n) = (\tilde{\mathbf{e}}_1 \ \cdots \ \tilde{\mathbf{e}}_n)P$, where $P \in \mathbb{F}^{n \times n}$, then assuming that $a_i = \boldsymbol{f}(\boldsymbol{e}_i)$ and $b_\ell = \boldsymbol{f}(\tilde{\mathbf{e}}_\ell)$ for $1 \leqslant i, \ell \leqslant n$, we have:

$$\begin{aligned} a_i = \boldsymbol{f}(\boldsymbol{e}_i) &= \boldsymbol{f}\left(\tilde{\mathbf{e}}_1 p_{1i} + \cdots + \tilde{\mathbf{e}}_n p_{ni}\right) \\ &= \sum_{\ell=1}^{n} \boldsymbol{f}(\tilde{\mathbf{e}}_\ell) p_{\ell i} = \sum_{\ell=1}^{n} b_\ell p_{\ell i}. \end{aligned}$$

Thus, the components of $\boldsymbol{f}$ relative to the two bases $(\boldsymbol{e}_1 \ \cdots \ \boldsymbol{e}_n)$, and $(\tilde{\mathbf{e}}_1 \ \cdots \ \tilde{\mathbf{e}}_n)$ transform in the same manner as these bases. We say that $\{a_1, \ldots, a_n\}$ are *covariant* components of the covector $\boldsymbol{f}$.

Let $C \in \mathbb{F}^{m \times n}$ be a matrix. If $\boldsymbol{x} \in \mathbb{F}^m$ and $\boldsymbol{y} \in \mathbb{F}^n$, the function $f_C : L \times M \longrightarrow \mathbb{F}$ defined by $f_C(\boldsymbol{x}, \boldsymbol{y}) = \boldsymbol{x}' C \boldsymbol{y}$ can be easily seen to be bilinear. The next theorem shows that all biliniar functions between two finite-dimensional spaces can be defined in this manner.

Theorem

Let L, M be two finite-dimensional $\mathbb{F}$-linear spaces. If $f : L \times M \longrightarrow \mathbb{F}$ is a bilinear form, then there is a matrix $C_f \in \mathbb{F}^{m \times n}$ such that

$$f(\boldsymbol{x}, \boldsymbol{y}) = \boldsymbol{x}' C_f \boldsymbol{y}$$

for all $\boldsymbol{x} \in L$ and $\boldsymbol{y} \in M$.

Proof

Suppose that $B = \{\boldsymbol{x}_1, \ldots, \boldsymbol{x}_m\}$ and $B' = \{\boldsymbol{y}_1, \ldots, \boldsymbol{y}_n\}$, are bases in L and M, respectively. Let $\boldsymbol{x} = a_1\boldsymbol{x}_1 + \cdots + a_m\boldsymbol{x}_m$ be the expression of $\boldsymbol{x} \in L$ in the base B. Similarly, let $\boldsymbol{y} = b_1\boldsymbol{y}_1 + \cdots + b_n\boldsymbol{y}_n$ be the expression of $\boldsymbol{y} \in M$ in B'. The bilinearily of f implies:

$$f(\boldsymbol{x}, \boldsymbol{y}) = f\left(\sum_{i=1}^{m} a_i\boldsymbol{x}_i, \sum_{j=1}^{n} b_j\boldsymbol{y}_j\right) = \sum_{i=1}^{m}\sum_{j=1}^{n} a_i b_j f(\boldsymbol{x}_i, \boldsymbol{y}_j).$$

If C_f is the matrix

$$C_f = \begin{pmatrix} f(\boldsymbol{x}_1, \boldsymbol{y}_1) & \cdots & f(\boldsymbol{x}_1, \boldsymbol{y}_n) \\ \vdots & \vdots & \vdots \\ f(\boldsymbol{x}_m, \boldsymbol{y}_1) & \cdots & f(\boldsymbol{x}_m, \boldsymbol{y}_n) \end{pmatrix},$$

then $f(\boldsymbol{x}, \boldsymbol{y}) = \boldsymbol{x}' C_f \boldsymbol{y}$, where

$$\boldsymbol{x} = \begin{pmatrix} a_1 \\ \vdots \\ a_m \end{pmatrix} \text{ and } \boldsymbol{y} = \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix}.$$

Let $A \in \mathbb{F}^{n \times n}$ be a symmetric matrix. The *quadratic form* associated to the matrix A is the function $f_A : \mathbb{R}^n \longrightarrow \mathbb{R}$ defined as $f_A(\boldsymbol{x}) = \boldsymbol{x}' A \boldsymbol{x}$ for $\boldsymbol{x} \in \mathbb{F}^n$. The *polar form* of the quadratic form f_A is the bilinear form $\tilde{f}_A$ defined by $\tilde{f}_A(\boldsymbol{x}, \boldsymbol{y}) = \boldsymbol{x}' A \boldsymbol{y}$ for $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{F}^n$.
Since $\boldsymbol{x}' A \boldsymbol{y}$ and $\boldsymbol{y}' A \boldsymbol{x}$ are scalars they are equal and we have:

$$\begin{aligned} f_A(\boldsymbol{x} + \boldsymbol{y}) &= (\boldsymbol{x} + \boldsymbol{y})' A (\boldsymbol{x} + \boldsymbol{y}) \\ &= \boldsymbol{x}' A \boldsymbol{x} + \boldsymbol{y}' A \boldsymbol{y} + \boldsymbol{x}' A \boldsymbol{y} + \boldsymbol{y}' A \boldsymbol{x} \\ &= f_A(\boldsymbol{x}) + f_A(\boldsymbol{y}) + 2 \tilde{f}_A(\boldsymbol{x}, \boldsymbol{y}), \end{aligned}$$

which allows us to express the polar form of f_A as

$$\tilde{f}_A(\boldsymbol{x}, \boldsymbol{y}) = \frac{1}{2}(f_A(\boldsymbol{x} + \boldsymbol{y}) - f_A(\boldsymbol{x}) - f_A(\boldsymbol{y})).$$

Let $h \in \text{Hom}(\mathbb{C}^m, \mathbb{C}^n)$ be a linear transformation between the linear spaces $\mathbb{C}^m$ and $\mathbb{C}^n$. Consider a basis in $\mathbb{C}^m$, $R = \{\boldsymbol{r}_1, \ldots, \boldsymbol{r}_m\}$, and a basis in $\mathbb{C}^n$, $S = \{\boldsymbol{s}_1, \ldots, \boldsymbol{s}_n\}$. The function h is completely determined by the images of the elements of the basis R, that is, by the set $\{h(\boldsymbol{r}_1), \ldots, h(\boldsymbol{r}_m)\}$. Indeed, if $\boldsymbol{x} = x_1\boldsymbol{r}_1 + \cdots + x_m\boldsymbol{r}_m$ and

$$h(\boldsymbol{r}_j) = a_{1j}\boldsymbol{s}_1 + a_{2j}\boldsymbol{s}_2 + \cdots + a_{nj}\boldsymbol{s}_n = \sum_{i=1}^{n} a_{ij}\boldsymbol{s}_i,$$

then, by linearity

$$\begin{aligned} h(\boldsymbol{x}) &= x_1 h(\boldsymbol{r}_1) + \cdots + x_m h(\boldsymbol{r}_m) \\ &= \sum_{j=1}^{m} x_j h(\boldsymbol{r}_j) \\ &= \sum_{j=1}^{m}\sum_{i=1}^{n} x_j a_{ij}\boldsymbol{s}_i. \end{aligned}$$

In a more compact form, we can write

$$h(\boldsymbol{x}) = (\boldsymbol{s}_1 \ \cdots \ \boldsymbol{s}_n) \begin{pmatrix} a_{11} & \cdots & a_{1m} \\ \vdots & \ldots & \vdots \\ a_{n1} & \cdots & a_{nm} \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_m \end{pmatrix}.$$

Let $\boldsymbol{x} \in \mathbb{C}^m$ be a vector such that $\boldsymbol{x} = x_1\boldsymbol{r}_1 + \cdots + x_m\boldsymbol{r}_m$. Then, the image of $\boldsymbol{x}$ under h is equals $A_h\boldsymbol{x}$, where A_h is

$$A_h = \begin{pmatrix} a_{11} & \cdots & a_{1m} \\ \vdots & \ldots & \vdots \\ a_{n1} & \cdots & a_{nm} \end{pmatrix}$$

Clearly, the matrix A_h attached to $h : \mathbb{C}^m \longrightarrow \mathbb{C}^n$ depends on the bases chosen for the linear spaces $\mathbb{C}^m$ and $\mathbb{C}^n$.

Let now $h : \mathbb{C}^n \longrightarrow \mathbb{C}^n$ be an endomorphism of $\mathbb{C}^n$ and let $R = \{\boldsymbol{r}_1, \ldots, \boldsymbol{r}_n\}$ and $S = \{\boldsymbol{s}_1, \ldots, \boldsymbol{s}_n\}$ be two bases of $\mathbb{C}^n$. The vectors $\boldsymbol{s}_i$ can be expressed as linear combinations of the vectors $\boldsymbol{r}_1, \ldots, \boldsymbol{r}_n$:

$$\boldsymbol{s}_i = p_{i1}\boldsymbol{r}_1 + \cdots + p_{in}\boldsymbol{r}_n, \tag{1}$$

for $1 \leqslant i \leqslant n$, which implies

$$h(\boldsymbol{s}_i) = p_{i1}h(\boldsymbol{r}_1) + \cdots + p_{in}h(\boldsymbol{r}_n). \tag{2}$$

for $1 \leqslant i \leqslant n$. Therefore, the matrix associated to a linear form $h : \mathbb{C}^m \longrightarrow \mathbb{C}$ is a column vector $\boldsymbol{r}$. In this case we can write $h(\boldsymbol{x}) = \boldsymbol{r}^{\mathsf{H}}\boldsymbol{x}$ for $\boldsymbol{x} \in \mathbb{R}^n$.

UMASS BOSTON

Theorem

Let $h \in Hom(\mathbb{C}^m, \mathbb{C}^n)$. The matrix $A_{h^*} \in \mathbb{C}^{m \times n}$ is the transposed of the matrix A_h, that is, we have $A_{h^*} = A'_h$.

By the previous discussion, if $\ell_1, \ldots, \ell_n$ is a basis of the space $(\mathbb{C}^n)^*$, then the j^{th} column of the matrix $A_{h^*} \in \mathbb{C}^{m \times n}$ is obtained by expressing the linear form $h^*(\ell_j) = \ell_j h$ in terms of a basis in the dual space $(\mathbb{C}^m)^*$. Therefore, we need to evaluate the linear form $\ell_j h \in (\mathbb{C}^m)^*$.

Let $\{\boldsymbol{p}_1, \ldots, \boldsymbol{p}_m\}$ be a basis in $\mathbb{C}^m$ and let $\{g_1, \ldots, g_m\}$ be its dual in $(\mathbb{C}^m)^*$. Also, let $\{\boldsymbol{q}_1, \ldots, \boldsymbol{q}_n\}$ be a basis in $\mathbb{C}^n$, and let $\{\ell_1, \ldots, \ell_n\}$ be its dual $(\mathbb{C}^n)^*$.

Observe that if $\boldsymbol{v} \in \mathbb{C}^m$ can be expressed as $\boldsymbol{v} = \sum_{j=1}^m v_j \boldsymbol{p}_j$, then

$$g_p(\boldsymbol{v}) = g_p\left(\sum_{j=1}^m v_j \boldsymbol{p}_j\right) = \sum_{j=1}^m v_j g_p(\boldsymbol{p}_j) = v_p,$$

because $\{g_1, \ldots, g_m\}$ is the dual of $\{\boldsymbol{p}_1, \ldots, \boldsymbol{p}_m\}$ in $(\mathbb{C}^m)^*$.

On the other hand, we can write

$$\begin{aligned}
\ell_j(h(\boldsymbol{v})) &= \ell_j\left(\sum_{p=1}^{m} v_p h(\boldsymbol{p}_p)\right) = \ell_j\left(\sum_{p=1}^{m} v_p \sum_{i=1}^{n} a_{ip}\boldsymbol{q}_i\right) \\
&= \ell_j\left(\sum_{p=1}^{m}\sum_{i=1}^{n} v_p a_{ip}\boldsymbol{q}_i\right) = \sum_{p=1}^{m}\sum_{i=1}^{n} v_p a_{ip}\ell_j(\boldsymbol{q}_i) \\
&= \sum_{p=1}^{m} v_p a_{jp} = \sum_{p=1}^{m} a_{jp} g_p(\boldsymbol{v}).
\end{aligned}$$

Thus, $h^*(\ell_j) = \sum_{p=1}^{m} a_{jp} g_p$ for every j, $1 \leq j \leqslant m$. This means that the j^{th} column of the matrix A_{h^*} is the transposed j^{th} row of the matrix A_h, so $A_{h^*} = (A_h)'$.

Matrix multiplication corresponds to the composition of linear mappings, as we show next.

Theorem

Let $h \in Hom(\mathbb{C}^m, \mathbb{C}^n)$ and $g \in Hom(\mathbb{C}^n, \mathbb{C}^p)$. Then,

$$A_{gh} = A_g A_h.$$

Proof.

If $\boldsymbol{p}_1, \ldots, \boldsymbol{p}_m$ is a basis for $\mathbb{C}^m$, then $A_{gh}(\boldsymbol{p}_i) = gh(\boldsymbol{p}_i) = g(h(\boldsymbol{p}_i)) = g(A_h \boldsymbol{p}_i) = A_g(A_h(\boldsymbol{p}_i))$ for every i, where $1 \leqslant i \leqslant n$. This proves that $A_{gh} = A_g A_h$. $\square$

The inverse direction, from matrices to linear operators is introduced next.

Definition

Let $A \in \mathbb{C}^{n\times m}$ be a matrix. The *linear operator associated to A*, is the mapping $h_A : \mathbb{C}^m \longrightarrow \mathbb{C}^n$ given by $h_A(\boldsymbol{x}) = A\boldsymbol{x}$ for $\boldsymbol{x} \in \mathbb{C}^m$.

If $\{\boldsymbol{e}_1, \ldots, \boldsymbol{e}_m\}$ is a basis for $\mathbb{C}^m$, then $h_A(\boldsymbol{e}_i)$ is the $i^{\rm th}$ column of the matrix A.

It is immediate that $A_{h_A} = A$ and $h_{A_h} = h$.

Attributes of a matrix A are usually transferred to the linear operator h_A.

For example, if A is Hermitian we say that h_A is Hermitian.

Attributes of a matrix A are usually transferred to the linear operator h_A. For example, if A is Hermitian we say that h_A is Hermitian.

Definition

Let $A \in \mathbb{C}^{n \times m}$ be a matrix. The *range of A* is the subspace $\mathrm{Im}(h_A)$ of $\mathbb{C}^n$. The *null space of A* is the subspace $\mathrm{Ker}(h_A)$.
The range of A and the null space of A are denoted by $\mathrm{range}(A)$ and $\mathrm{null}(A)$, respectively.

Clearly, $C_{A,n} = \mathrm{range}(A)$. The null space of $A \in \mathbb{C}^{m \times n}$ consists of those $\boldsymbol{x} \in \mathbb{C}^n$ such that $A\boldsymbol{x} = \mathbf{0}$.

Let $\{\boldsymbol{p}_1, \ldots, \boldsymbol{p}_m\}$ be a basis of $\mathbb{C}^m$. Since $\text{range}(A) = \text{Im}(h_A)$ it follows that this subspace is generated by the set $\{h_A(\boldsymbol{p}_1), \ldots, h_A(\boldsymbol{p}_m)\}$, that is, by the columns of the matrix A. For this reason the subspace $\text{range}(A)$ is also known as the *column subspace* of A.

Theorem

Let $A, B \in \mathbb{C}^{m \times n}$ be two matrices. Then

$$\textit{range}(A + B) \subseteq \textit{range}(A) + \textit{range}(B).$$

Proof.

Let $u \in \text{range}(A + B)$. There exists $v \in \mathbb{C}^n$ such that $u = (A + B)v = Av + Bv$. If $x = Av$ and $y = Bv$, we have $x \in \text{range}(A)$ and $y \in \text{range}(B)$, so $u = x + y \in \text{range}(A) + \text{range}(B)$. □

Definition

The *rank* of a matrix A is the number denoted by $\text{rank}(A)$ given by $\text{rank}(A) = \dim(\text{range}(A)) = \dim(\text{Im}(h_A))$.

Thus, *the rank of A is the maximal size of a set of linearly independent columns of A.*
A previous theorem applied to the linear mapping $h_A : \mathbb{C}^m \longrightarrow \mathbb{C}^n$ means that for $A \in \mathbb{C}^{n\times m}$ we have:

$$\dim(\text{null}(A)) + \text{rank}(A) = m. \tag{3}$$

Observe that if $A \in \mathbb{C}^{m\times m}$ is non-singular, then $A\boldsymbol{x} = \mathbf{0}_m$ implies $\boldsymbol{x} = \mathbf{0}_m$. Thus, if $\boldsymbol{x} \in \text{null}(A) \cap \text{range}(A)$ it follows that $A\boldsymbol{x} = \mathbf{0}$, so the subspaces $\text{null}(A)$ and $\text{range}(A)$ are complementary.

Example

For the matrix

$$A = \begin{pmatrix} 1 & 0 & 2 \\ 1 & -1 & 1 \\ 2 & 1 & 5 \\ 1 & 2 & 4 \end{pmatrix}$$

we have $\text{rank}(A) = 2$. Indeed, if $\mathbf{c}_1, \mathbf{c}_2, \mathbf{c}_3$ are its columns, then it is easy to see that $\{\mathbf{c}_1, \mathbf{c}_2\}$ is a linearly independent set, and $\mathbf{c}_3 = 2\mathbf{c}_1 + \mathbf{c}_2$. Thus, the maximal size of a set of linearly independent columns of A is 2.

Example cont'd

The number of linearly independent rows of

$$A = \begin{pmatrix} 1 & 0 & 2 \\ 1 & -1 & 1 \\ 2 & 1 & 5 \\ 1 & 2 & 4 \end{pmatrix}$$

is also 2. Indeed, we have

$$(2\ 1\ 5) = a(1\ 0\ 2) + b(1\ \ -1\ 1)$$

for $a = 3$ and $b = -1$. Also,

$$(1\ 2\ 4) = c(1\ 0\ 2) + d(1\ \ -1\ 1)$$

for $c = 3$ and $d = -2$.

Example

Let $A \in \mathbb{C}^{n \times m}$ and $B \in \mathbb{C}^{p \times q}$. For the matrix $C \in \mathbb{C}^{(n+p)\times(m+q)}$ defined by

$$C = \begin{pmatrix} A & O_{n,q} \\ O_{p,m} & B \end{pmatrix}$$

we have $\text{rank}(C) = \text{rank}(A) + \text{rank}(B)$.
Suppose that $\text{rank}(C) = \ell$ and let $\boldsymbol{c}_1, \ldots, \boldsymbol{c}_\ell$ be a maximal set of linearly independent columns of C. Without loss of generality we may assume that the first k columns are among the first m columns of A and the remaining $\ell - k$ columns are among the last q columns of C. The first k columns of C correspond to k linearly independent columns of A, while the last $\ell - k$ columns correspond to $\ell - k$ linearly independent columns of B. Thus, $\text{rank}(C) = k \leqslant \text{rank}(A) + \text{rank}(B)$.

UMASS BOSTON

Example cont'd

Example

Conversely, suppose that $\text{rank}(A) = s$ and $\text{rank}(B) = t$. Let $\boldsymbol{a}_{i_1}, \ldots, \boldsymbol{a}_{i_s}$ be a maximal set of linearly independent columns of A and let $\boldsymbol{b}_{j_1}, \ldots, \boldsymbol{b}_{j_t}$ be a maximal set of linearly independent columns of B. Then, it is easy to see that the vectors

$$\begin{pmatrix} \boldsymbol{a}_{i_1} \\ \boldsymbol{0}_n \end{pmatrix}, \ldots, \begin{pmatrix} \boldsymbol{a}_{i_s} \\ \boldsymbol{0}_n \end{pmatrix}, \ldots, \begin{pmatrix} \boldsymbol{0}_n \\ \boldsymbol{b}_{j_1} \end{pmatrix}, \ldots, \begin{pmatrix} \boldsymbol{0}_n \\ \boldsymbol{b}_{j_t} \end{pmatrix}$$

constitute a linearly independent set of columns of C, so $\text{rank}(A) + \text{rank}(B) \leqslant \text{rank}(C)$. Thus, $\text{rank}(C) = \text{rank}(A) + \text{rank}(B)$.

Example

Let $\boldsymbol{x}$ and $\boldsymbol{y}$ be two vectors in $\mathbb{C}^n - \{\mathbf{0}\}$. The matrix $\boldsymbol{x}\boldsymbol{y}^{\mathsf{H}}$ has rank 1. Indeed, if $\boldsymbol{y}^{\mathsf{H}} = (y_1, y_2, \ldots, y_n)$, then we can write

$$\boldsymbol{x}\boldsymbol{y}^{\mathsf{H}} = (y_1\boldsymbol{x}\ y_2\boldsymbol{x}\ \cdots\ y_n\boldsymbol{x}),$$

which implies that the maximum number of linearly independent columns of $\boldsymbol{x}\boldsymbol{y}^{\mathsf{H}}$ is 1.

Example

Let $A, B \in \mathbb{C}^{n \times m}$. We have $\text{rank}(A + B) \leqslant \text{rank}(A) + \text{rank}(B)$.
Let $A = (\boldsymbol{a}_1 \ \boldsymbol{a}_2 \ \cdots \ \boldsymbol{a}_m)$ and $B = (\boldsymbol{b}_1 \ \boldsymbol{b}_2 \ \cdots \ \boldsymbol{b}_m)$ be two matrices, where $\boldsymbol{a}_1, \ldots, \boldsymbol{a}_m, \boldsymbol{b}_1, \ldots, \boldsymbol{b}_m \in \mathbb{C}^n$. Clearly, we have

$$A + B = (\boldsymbol{a}_1 + \boldsymbol{b}_1 \ \boldsymbol{a}_2 + \boldsymbol{b}_2 \ \cdots \ \boldsymbol{a}_m + \boldsymbol{b}_m).$$

If $\boldsymbol{x} \in \text{Im}(A + B)$ we can write:

$$\boldsymbol{x} = x_1(\boldsymbol{a}_1 + \boldsymbol{b}_1) + x_2(\boldsymbol{a}_2 + \boldsymbol{b}_2) + \cdots + x_m(\boldsymbol{a}_m + \boldsymbol{b}_m) = \boldsymbol{y} + \boldsymbol{z},$$

where

$$\begin{aligned} \boldsymbol{y} &= x_1\boldsymbol{a}_1 + \cdots + x_m\boldsymbol{a}_m \in \text{Im}(A), \\ \boldsymbol{z} &= x_1\boldsymbol{b}_1 + \cdots + x_m\boldsymbol{b}_m \in \text{Im}(B). \end{aligned}$$

Thus, $\text{Im}(A + B) \subseteq \text{Im}(A) + \text{Im}(B)$. Since the dimension of the sum of two subspaces of a linear space is less or equal to the dimension of sum of these subspaces, the result follows.

The above discussion also shows that if $A \in \mathbb{C}^{n\times m}$, then $\text{rank}(A) \leqslant \min\{m, n\}$.

Theorem

Let $A \in \mathbb{C}^{m\times n}$ be a matrix. We have $\text{rank}(A) = \text{rank}(\overline{A})$.

Proof.

Suppose that $A = (\mathbf{a}_1, \ldots, \mathbf{a}_n)$ and that the set $\{\mathbf{a}_{i_1}, \ldots, \mathbf{a}_{i_p}\}$ is a set of linearly independent columns of A. Then, the set $\{\overline{\mathbf{a}_{i_1}}, \ldots, \overline{\mathbf{a}_{i_p}}\}$ is a set of linearly independent columns of $\overline{A}$. This implies $\text{rank}(\overline{A}) = \text{rank}(A)$. □

Corollary

We have $rank(A) = rank(A^H)$ *for every matrix* $A \in \mathbb{C}^{m \times n}$.

Proof.

Since $A^{\mathrm{H}} = \overline{A'}$, the statement follows immediately. □

If $A \in \mathbb{C}^{m \times n}$ is a full-rank matrix and $m \geqslant n$, then the n columns of the matrix are linearly independent; similarly, if $n \geqslant m$, the m rows of the matrix are linearly independent.

A matrix that is not a full-rank is said to be *degenerate*. A degenerate square matrix is said to be *singular*. A *non-singular matrix* $A \in \mathbb{C}^{n \times n}$ is a matrix that is not singular and, therefore has $\text{rank}(A) = n$.

Theorem

A matrix $A \in \mathbb{C}^{n \times n}$ is non-singular if and only if it is invertible.

Proof

Suppose that A is non-singular, that is, $\text{rank}(A) = n$. In other words the set of columns $\{\mathbf{c}_1, \ldots, \mathbf{c}_n\}$ of A is linearly independent, and therefore, is a basis of $\mathbb{C}^n$. Then, each of the vectors $\mathbf{e}_i$ can be expressed as a unique combination of the columns of A, that is

$$\mathbf{e}_i = b_{1i}\mathbf{c}_1 + b_{2i}\mathbf{c}_2 + \cdots + b_{ni}\mathbf{c}_n,$$

for $1 \leqslant i \leqslant n$. These equalities can be written as

$$(\mathbf{c}_1 \ \cdots \ \mathbf{c}_n) \begin{pmatrix} b_{11} & \cdots & b_{1n} \\ b_{21} & \cdots & b_{2n} \\ \vdots & \cdots & \vdots \\ b_{n1} & \cdots & b_{nn} \end{pmatrix} = I_n.$$

Consequently, the matrix A is invertible and

$$A^{-1} = \begin{pmatrix} b_{11} & \cdots & b_{1n} \\ b_{21} & \cdots & b_{2n} \\ \vdots & \cdots & \vdots \\ b_{n1} & \cdots & b_{nn} \end{pmatrix}$$

Proof cont'd

Suppose now that A is invertible and that

$$d_1\mathbf{c}_1 + \cdots + d_n\mathbf{c}_n = \mathbf{0}.$$

This is equivalent to

$$A \begin{pmatrix} d_1 \\ \vdots \\ d_n \end{pmatrix} = \mathbf{0}.$$

Multiplying both sides by A^{-1} implies

$$\begin{pmatrix} d_1 \\ \vdots \\ d_n \end{pmatrix} = \mathbf{0},$$

so $d_1 = \cdots = d_n = 0$, which means that the set of columns of A is linearly independent, so $\text{rank}(A) = n$.

Corollary

A matrix $A \in \mathbb{C}^{n \times n}$ *is non-singular if and only if* $A\mathbf{x} = \mathbf{0}$ *implies* $\mathbf{x} = \mathbf{0}$ *for* $\mathbf{x} \in \mathbb{C}^n$.

Proof.

If A is non-singular then A is invertible. Therefore, $A\mathbf{x} = \mathbf{0}$ implies $A^{-1}(Ax) = A^{-1}\mathbf{0}$, so $\mathbf{x} = \mathbf{0}$.
Conversely, suppose that $A\mathbf{x} = \mathbf{0}$ implies $\mathbf{x} = \mathbf{0}$. If $A = (\mathbf{c}_1 \ \cdots \ \mathbf{c}_n)$ and $\mathbf{x} = (x_1, \ldots, x_n)'$, the previous implication means that $x_1\mathbf{c}_1 + \cdots + x_n\mathbf{c}_n = \mathbf{0}$ implies $x_1 = \cdots = x_n = 0$, so $\{\mathbf{c}_1, \ldots, \mathbf{c}_n\}$ is linearly independent. Therefore, $\text{rank}(A) = n$, so A is non-singular. □

Define the *similarity* relation "$\sim$" on the set of square matrices $\mathbb{C}^{n\times n}$ by $A \sim B$ if there exists an invertible matrix X such that $A = XBX^{-1}$.
If X is a unitary matrix, then we say that A and B are *unitarily similar* and we write $A \sim_u B$, so $\sim_u$ is a subset of $\sim$. In this case, we have $A = XBX^{\mathsf{H}}$. o

Theorem

The relations "$\sim$" and "$\sim_u$" are equivalence relations.

Proof.

We have $A \sim A$ because $A = I_n A (I_n)^{-1}$, so $\sim$ is a reflexive relation. To prove that $\sim$ is symmetric suppose that $A = XBX^{-1}$. Then, $B = X^{-1}AX$ and, since X^{-1} is invertible, we have $B \sim A$.
Finally, to verify the transitivity, let A, B, C be such that $A = XBX^{-1}$ and $B = YCY^{-1}$, where X and Y are two invertible matrices. This allows us to write

$$A = XBX^{-1} = XYCY^{-1}X^{-1} = (XY)C(XY)^{-1},$$

which proves that $A \sim C$.
We leave to the reader the similar proof concerning $\sim_u$. □

Theorem

If $A \sim_u B$, where $A, B \in \mathbb{C}^{n \times n}$, then $A^H A \sim_u B^H B$.

Proof.

Since $A \sim_u B$ there exists a unitary matrix X such that $A = XBX^{-1} = XBX^{\mathrm{H}}$. Then, $A^{\mathrm{H}} = XB^{\mathrm{H}}X^{\mathrm{H}}$, so $A^{\mathrm{H}}A = XB^{\mathrm{H}}X^{\mathrm{H}}XBX^{\mathrm{H}} = XB^{\mathrm{H}}BX^{\mathrm{H}}$. Thus, $A^{\mathrm{H}}A$ is unitarily similar to $B^{\mathrm{H}}B$. □

Theorem

Let A and B be two matrices in $\mathbb{C}^{m\times n}$*. We have* $A \sim B$ *if and only if* $rank(A) = rank(B)$*.*

Proof: If $A \in \mathbb{C}^{m\times n}$ be a matrix with $\text{rank}(A) = r > 0$, then

$$A \sim \begin{pmatrix} I_r & O_{r,n-r} \\ O_{m-r,r} & O_{m-r,n-r} \end{pmatrix}.$$

Thus, for every two matrices $A, B \in \mathbb{C}^{n\times m}$ of rank r we have $A \sim B$ because both are similar to

$$\begin{pmatrix} I_r & O_{r,n-r} \\ O_{m-r,r} & O_{m-r,n-r} \end{pmatrix}.$$

Proof cont'd

Conversely, suppose that $A \sim B$, that is, $A = GBH$, where $G \in \mathbb{C}^{m\times m}$ and $H \in \mathbb{C}^{n\times n}$ are non-singular matrices. By a previous corollary we have $\text{rank}(A) = \text{rank}(B)$.

Definition

A matrix $A \in \mathbb{C}^{n \times n}$ is *diagonalizable* if there exists a diagonal matrix D such that $A \sim D$.
Let $\mathcal{M}$ be a class of matrices. A is $\mathcal{M}$-*diagonalizable* if there exists a matrix $M \in \mathcal{M}$ such that $A = MDM^{-1}$.

For example, if A is $\mathcal{M}$-diagonalizable and $\mathcal{M}$ is the class of unitary matrices we say that A is *unitarily diagonalizable*.

Let $f : \mathbb{C}^n \longrightarrow \mathbb{C}$ be a polynomial given by

$$f(z) = a_0 z^n + a_1 z^{n-1} + \cdots + a_n,$$

where $a_0, a_1, \ldots, a_n \in \mathbb{C}$. If $A \in \mathbb{C}^{m \times m}$, then the matrix $f(A)$ is defined by

$$f(A) = a_0 A^n + a_1 A^{n-1} + \cdots + a_n I_m.$$

Theorem

If $T \in \mathbb{C}^{m \times m}$ is an upper (a lower) triangular matrix and f is a polynomial, then $f(T)$ is an upper (a lower) triangular matrix. Furthermore, if the diagonal elements of T are $t_{11}, t_{22}, \ldots, t_{mm}$, then the diagonal elements of $f(T)$ are $f(t_{11}), f(t_{22}), \ldots, f(t_{mm})$, respectively.

Proof

Any power T^k of T is an upper (a lower) triangular matrix. Since the sum of upper (lower) triangular matrices is upper (lower) triangular, if follows that $f(T)$ is an upper triangular (a lower triangular) matrix.
An easy argument by induction on k (left to the reader) shows that if the diagonal elements of T are $t_{11}, t_{22}, \ldots, t_{mm}$, then the diagonal elements of T^k are $t_{11}^k, t_{22}^k, \ldots, t_{mm}^k$. The second part of the theorem follows immediately.

Theorem

Let $A, B \in \mathbb{C}^{m\times m}$. If $A \sim B$ and f is a polynomial, then $f(A) \sim f(B)$.

Proof.

Let X be an invertible matrix such that $A = XBX^{-1}$. It is straightforward to verify that $A^k = XB^kX^{-1}$ for $k \in \mathbb{N}$. This implies that $f(A) = Xf(B)X^{-1}$, so $f(A) \sim f(B)$. then $f(A) \sim f(B)$. □

UMASS BOSTON

Definition

Let A and B be two matrices in $\mathbb{C}^{n\times n}$. The matrices A and B are *congruent* if there exists an invertible matrix $X \in \mathbb{C}^{n\times n}$ such that $B = XAX^{\mathsf{H}}$. This is denoted by $A \sim_{\mathsf{H}} B$.

The relation $\sim_{\mathsf{H}}$ is an equivalence on $\mathbb{C}^{n\times n}$. We have $A \sim_{\mathsf{H}} A$ because $A = I_n A I_n^{\mathsf{H}}$. If $A \sim_{\mathsf{H}} B$, then $B = XAX^{\mathsf{H}}$, so $A = X^{-1}B(X^{\mathsf{H}})^{-1} = X^{-1}B(X^{-1})^{\mathsf{H}}$, which implies $B \sim_{\mathsf{H}} A$. Finally, $\sim_{\mathsf{H}}$ is transitive because if $B = XAX^{\mathsf{H}}$ and $C = YBY^{\mathsf{H}}$, where X and Y are invertible matrices, then $C = (YX)A(YX)^{\mathsf{H}}$ and YX is an invertible matrix. It is immediate that any two congruent matrices have the same rank.

Recapitulation

- A and B are *similar matrices*, $A \sim B$, if there exists an invertible matrix X such that $A = XBX^{-1}$;
- A and B are *congruent matrices*, $A \sim_{\mathrm{H}} B$, if there exists an invertible matrix $X \in \mathbb{C}^{n \times n}$ such that $B = XAX^{\mathrm{H}}$;
- A and B are *unitarily similar*, $A \sim_u B$, if there exists a unitary matrix U such that $A = UBU^{-1}$.

Since every unitary matrix is invertible and its inverse equals its conjugate Hermitian matrix, it follows that $\sim_u$ is a subset of both $\sim$ and $\sim_{\mathrm{H}}$.

Definition

Let $A \in \mathbb{C}^{m\times n}$ and $B \in \mathbb{C}^{p\times q}$ be two matrices. The *Kronecker product* of these matrices is the matrix $A \otimes B \in \mathbb{C}^{mp\times nq}$ defined by

$$A \otimes B = \begin{pmatrix} a_{11}B & a_{12}B & \cdots & a_{1n}B \\ a_{21}B & a_{22}B & \cdots & a_{2n}B \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1}B & a_{m2}B & \cdots & a_{mn}B \end{pmatrix}.$$

The Kronecker product $A \otimes B$ creates mn copies of the matrix B and multiplies each copy by the corresponding element of A.

Consider the matrices

$$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \text{ and } B = \begin{pmatrix} b_{11} & b_{12} & b_{13} \\ b_{21} & b_{22} & b_{23} \\ b_{31} & b_{32} & b_{33} \end{pmatrix}$$

Their Kronecker product is

$$A \otimes B = \left(\begin{array}{ccc|ccc} a_{11}b_{11} & a_{11}b_{12} & a_{11}b_{13} & a_{12}b_{11} & a_{12}b_{12} & a_{12}b_{13} \\ a_{11}b_{21} & a_{11}b_{22} & a_{11}b_{23} & a_{12}b_{21} & a_{12}b_{22} & a_{12}b_{23} \\ a_{11}b_{31} & a_{11}b_{32} & a_{11}b_{33} & a_{12}b_{31} & a_{12}b_{32} & a_{12}b_{33} \\ \hline a_{21}b_{11} & a_{21}b_{12} & a_{21}b_{13} & a_{22}b_{11} & a_{22}b_{12} & a_{22}b_{13} \\ a_{21}b_{21} & a_{21}b_{22} & a_{21}b_{23} & a_{22}b_{21} & a_{22}b_{22} & a_{22}b_{23} \\ a_{21}b_{31} & a_{21}b_{32} & a_{21}b_{33} & a_{22}b_{31} & a_{22}b_{32} & a_{22}b_{33} \end{array}\right).$$

Let $C \in \mathbb{C}^{mp \times nq}$ be the Kronecker product of the matrices $A \in \mathbb{C}^{m \times n}$ and $B \in \mathbb{C}^{p \times q}$. We seek to express the value of c_{ij}, where $1 \leqslant i \leqslant mp$ and $1 \leqslant j \leq nq$. It is easy to see that

$$c_{ij} = a_{\lceil \frac{i}{p} \rceil, \lceil \frac{j}{q} \rceil} b_{i-p\left(\lceil \frac{i}{p} \rceil - 1\right), j-q\left(\lceil \frac{j}{q} \rceil - 1\right)}. \tag{4}$$

Theorem

For any matrices A, B, C, D *we have:*

- $(A \otimes B)' = A' \otimes B'$,
- $(A \otimes B) \otimes C = A \otimes (B \otimes C)$,
- $(A \otimes B)(C \otimes D) = (AC \otimes BD)$,
- $A \otimes B + A \otimes C = A \otimes (B + C)$,
- $A \otimes D + B \otimes D = (A + B) \otimes D$,
- $(A \otimes B)' = A' \otimes B'$,
- $(A \otimes B)^H = A^H \otimes B^H$,

when the usual matrix sum and multiplication are well-defined in each of the above equalities.

Example

Let $\boldsymbol{x} \in \mathbb{C}^n$ and $\boldsymbol{y} \in \mathbb{C}^m$. We have

$$\boldsymbol{x} \otimes \boldsymbol{y} = \begin{pmatrix} x_1 \boldsymbol{y} \\ \vdots \\ x_n \boldsymbol{y} \end{pmatrix} = \begin{pmatrix} y_1 \boldsymbol{x} \\ \vdots \\ y_m \boldsymbol{x} \end{pmatrix} \in \mathbb{C}^{mn}.$$

Theorem

If $A \in \mathbb{C}^{n\times n}$ and $B \in \mathbb{C}^{m\times m}$ are two invertible matrices, then $A \otimes B$ is invertible and $(A \otimes B)^{-1} = A^{-1} \otimes B^{-1}$.

Proof.

Since

$$(A \otimes B)(A^{-1} \otimes B^{-1}) = (AA^{-1} \otimes BB^{-1}) = I_n \otimes I_m,$$

the theorem follows by noting that $I_n \otimes I_m = I_{nm}$. □

Theorem

Let $A \in \mathbb{C}^{n \times n}$ and $B \in \mathbb{C}^{m \times m}$ be two normal (unitary) matrices. Their Kronecker product $A \otimes B$ is also a normal (a unitary) matrix.

Proof.

We can write

$$\begin{aligned} (A \otimes B)'(A \otimes B) &= (A' \otimes B')(A \otimes B) \\ &= (A'A \otimes B'B) \\ &= (AA' \otimes BB') \\ &\quad \text{(because both } A \text{ and } B \text{ are normal)} \\ &= (A \otimes B)(A \otimes B)', \end{aligned}$$

which implies that $A \otimes B$ is normal. □

UMASS BOSTON

Definition

Let $A \in \mathbb{C}^{m\times m}$ and $B \in \mathbb{C}^{n\times n}$ be two square matrices. Their *Kronecker sum* is the matrix $A \oplus B \in \mathbb{C}^{mn\times mn}$ defined by

$$A \oplus B = (A \otimes I_n) + (I_m \otimes B).$$

The *Kronecker difference* is the matrix $A \ominus B \in \mathbb{C}^{mn\times mn}$ defined by

$$A \ominus B = (A \otimes I_n) - (I_m \otimes B).$$

Definition

Let $A, B \in \mathbb{C}^{m \times n}$. The *Hadamard product* of A and B is the matrix $A \odot B \in \mathbb{C}^{m \times n}$ defined by

$$A \odot B = \begin{pmatrix} a_{11}b_{11} & a_{12}b_{12} & \cdots & a_{1n}b_{1n} \\ a_{21}b_{21} & a_{22}b_{22} & \cdots & a_{2n}b_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1}b_{m1} & a_{m2}b_{m2} & \cdots & a_{mn}b_{mn} \end{pmatrix}.$$

Definition

The *Hadamard quotient* $A \oslash B$ is defined only if $b_{ij} \neq 0$ for $1 \leqslant i \leqslant m$ and $1 \leqslant j \leqslant n$. In this case

$$A \oslash B = \begin{pmatrix} \frac{a_{11}}{b_{11}} & \frac{a_{12}}{b_{12}} & \cdots & \frac{a_{1n}}{b_{1n}} \\ \frac{a_{21}}{b_{21}} & \frac{a_{22}}{b_{22}} & \cdots & \frac{a_{2n}}{b_{2n}} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{a_{m1}}{b_{m1}} & \frac{a_{m2}}{b_{m2}} & \cdots & \frac{a_{mn}}{b_{mn}} \end{pmatrix}.$$

Theorem

If $A, B, C \in \mathbb{C}^{m\times n}$ *and* $c \in \mathbb{C}$ *we have*

- $A \odot B = B \odot A$;
- $A \odot J_{m,n} = J_{m,n} \odot A = A$;
- $A \odot (B + C) = A \odot B + A \odot C$;
- $A \odot (cB) = c(A \odot B)$.

Note that the Hadamard product of two matrices $A, B \in \mathbb{C}^{m \times n}$ is a submatrix of the Kronecker product $A \otimes B$.

Example

Let $A, B \in \mathbb{C}^{2 \times 3}$ be the matrices

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{pmatrix} \text{ and } B = \begin{pmatrix} b_{11} & b_{12} & b_{13} \\ b_{21} & b_{22} & b_{23} \end{pmatrix}.$$

The Kronecker product of these matrices is $A \otimes B \in \mathbb{C}^{4 \times 9}$ given by:

$$A \otimes B = \begin{pmatrix} a_{11} b_{11} & a_{11} b_{12} & a_{11} b_{13} & a_{12} b_{11} & a_{12} b_{12} & a_{12} b_{13} & a_{13} b_{11} & a_{13} b_{12} \\ a_{11} b_{21} & a_{11} b_{22} & a_{11} b_{23} & a_{12} b_{21} & a_{12} b_{22} & a_{12} b_{23} & a_{13} b_{21} & a_{13} b_{22} \\ a_{21} b_{11} & a_{21} b_{12} & a_{21} b_{13} & a_{22} b_{11} & a_{22} b_{12} & a_{22} b_{13} & a_{23} b_{11} & a_{23} b_{12} \\ a_{21} b_{21} & a_{21} b_{22} & a_{21} b_{23} & a_{22} b_{21} & a_{22} b_{22} & a_{22} b_{23} & a_{23} b_{21} & a_{23} b_{22} \end{pmatrix}$$

Example

The Hadamard product of the same matrices is

$$A \odot B = \begin{pmatrix} a_{11}b_{11} & a_{12}b_{12} & a_{13}b_{13} \\ a_{21}b_{21} & a_{22}b_{22} & a_{23}b_{23} \end{pmatrix},$$

and we can regard the Hadamard product as a submatrix of the Kronecker product $A \otimes B$,

$$A \odot B = (A \otimes B) \begin{bmatrix} 1,5,9 \\ 4,4,4 \end{bmatrix}.$$

Another matrix product involves matrices that have the same number of columns.

Definition

Let $A \in \mathbb{C}^{m \times n}$ and $B \in \mathbb{C}^{p \times n}$ be two matrices that have n columns,

$$A = (\boldsymbol{a}_1 \ \cdots \ \boldsymbol{a}_n) \text{ and } B = (\boldsymbol{b}_1 \ \cdots \ \boldsymbol{b}_n).$$

The *Khatri-Rao* product of A and B is the matrix

$$A * B = (\boldsymbol{a}_1 \otimes \boldsymbol{b}_1 \ \boldsymbol{a}_2 \otimes \boldsymbol{b}_2 \ \cdots \ \boldsymbol{a}_n \otimes \boldsymbol{b}_n).$$

Example

The Khatri-Rao product of the matrices

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix} \text{ and} B = \begin{pmatrix} 1 & 0 & 2 \\ 2 & 1 & 3 \\ -1 & 2 & 1 \end{pmatrix}$$

is the matrix $(\boldsymbol{a}_1 \otimes \boldsymbol{b}_1 \ \boldsymbol{a}_2 \otimes \boldsymbol{b}_2 \ \boldsymbol{a}_3 \otimes \boldsymbol{b}_3)$ which equals

$$\begin{pmatrix} 1 & 0 & 6 \\ 2 & 2 & 9 \\ -1 & 4 & 3 \\ 4 & 0 & 12 \\ 8 & 5 & 18 \\ -4 & 10 & 6 \end{pmatrix}.$$

Definition

Let $\boldsymbol{u} \in \mathbb{C}^m$ and $\boldsymbol{v} \in \mathbb{C}^n$. The *outer product* of of the vectors $\boldsymbol{u}$ and $\boldsymbol{v}$ is the matrix $\boldsymbol{u} * \boldsymbol{v} \in \mathbb{C}^{m \times n}$ defined by $\boldsymbol{u} * \boldsymbol{v} = \boldsymbol{u}\boldsymbol{v}^{\mathsf{H}}$.

The outer product of two vectors is a matrix of rank 1.
For $\boldsymbol{u} \in \mathbb{C}^m$ and $\boldsymbol{v} \in \mathbb{C}^n$ we have $\boldsymbol{v} * \boldsymbol{u} = \boldsymbol{v}\boldsymbol{u}^{\mathsf{H}} = (\boldsymbol{u}\boldsymbol{v}^{sH})^{\mathsf{H}} = (\boldsymbol{u} * \boldsymbol{v})^{\mathsf{H}}$.
Therefore, the outer product is not commutative because for $\boldsymbol{u} \in \mathbb{C}^m$ and $\boldsymbol{v} \in \mathbb{C}^n$ we have $\boldsymbol{u} * \boldsymbol{v} \in \mathbb{C}^{m \times n}$ and $\boldsymbol{v} * \boldsymbol{u} \in \mathbb{C}^{n \times m}$.
Note that when $m = n$ we have $\boldsymbol{u}\boldsymbol{v}^{\mathsf{H}} = \text{trace}(\boldsymbol{u} * \boldsymbol{v})$.

UMASS BOSTON

Example

Let

$$\boldsymbol{u} = \begin{pmatrix} u_1 \\ u_2 \\ u_3 \end{pmatrix} \text{ and } \boldsymbol{v} = \begin{pmatrix} v_1 \\ v_2 \end{pmatrix}.$$

We have

$$\boldsymbol{u} * \boldsymbol{v} = \begin{pmatrix} u_1 v_1 & u_1 v_2 \\ u_2 v_1 & u_2 v_2 \\ u_3 v_1 & u_3 v_2 \end{pmatrix} \text{ and } \boldsymbol{v} * \boldsymbol{u} = \begin{pmatrix} v_1 u_1 & v_1 u_2 & v_1 u_3 \\ v_2 u_1 & v_2 u_2 & v_2 u_3 \end{pmatrix}.$$

Example

Contrast this with the Kronecker products:

$$\boldsymbol{u} \otimes \boldsymbol{v} = \begin{pmatrix} u_1 v_1 \\ u_1 v_2 \\ u_2 v_1 \\ u_2 v_2 \\ u_3 v_1 \\ u_3 v_2 \end{pmatrix} \text{ and } \boldsymbol{v} \otimes \boldsymbol{u} = \begin{pmatrix} v_1 u_1 \\ v_1 u_2 \\ v_1 u_3 \\ v_2 u_1 \\ v_2 u_2 \\ v_2 u_3 \end{pmatrix}.$$

Note that the entries of the Kronecker product $\boldsymbol{u} \otimes \boldsymbol{v}$ can be obtained by reading the entries of $\boldsymbol{u} * \boldsymbol{v}$ row-wise, or the entries of the same column-wise. Similar statements hold for $\boldsymbol{v} \otimes \boldsymbol{u}$. This observation suggested the use of the Kronecker symbol $\otimes$ for outer products of vectors. In other words, we will denote the outer products $\boldsymbol{u} * \boldsymbol{v}$ and $\boldsymbol{v} * \boldsymbol{u}$ with $\boldsymbol{u} \otimes \boldsymbol{v}$ and $\boldsymbol{v} \otimes \boldsymbol{u}$, respectively.