MA580H Matrix Computations

Lectures 1 & 2: Vectors and Matrices

Rafikul Alam
Department of Mathematics
IIT Guwahati

Outline

Topics:

- Vectors in $\mathbb{R}^n$ and $\mathbb{C}^n$
- Matrix-vector multiplication
- Matrix-matrix multiplication
- Block matrices
- Outer product of vectors

Course Syllabus

Linear systems: All variants of Gaussian elimination and LU factorization, Cholesky factorization.

Linear least-squares problem: Normal equations, rotators and reflectors, QR factorization via rotators, reflectors and Gram Schmidt orthonormalisation, QR method for linear least-squares problems, rank deficient least-squares problems.

Singular value decomposition (SVD): Numerical rank determination via SVD, solution of least squares problems, Moore- Penrose inverse, low rank approximations via SVD, Principal Component Analysis, applications to data mining and image recognition.

Eigenvalue Decomposition: Power, inverse power and Rayleigh quotient iterations, Schur's decomposition, unitary similarity transformation of Hermitian matrices to tridiagonal form, QR algorithm, implementation of explicit QR algorithm for Hermitian matrices.

Textbooks

- L. N. Trefethen and David Bau, Numerical Linear Algebra, SIAM, Philadelphia, 1997.
- D. S. Watkins, Fundamentals of Matrix Computations, 2nd Edition, Wiley, 2002.
- L. Elden, Matrix Methods in Data Mining and Pattern Recognition, SIAM, Philadelphia, 2007.

Another good book on Least-Squares problems:

- S. Boyd and L. Vandenberghe, Introduction to Applied Linear Algebra: Vectors, Matrices and Least Squares, Cambridge University Press, 2018

Vectors in $\mathbb{R}^n$

We define $\mathbb{R}^n$ to be the set of all ordered n-tuples of real numbers. Thus an n-tuple in $\mathbb{R}^n$ (also called an n-vector) is of the form

$$\text{row vector:} \quad \mathbf{v} = [v_1, \dots, v_n] \quad \text{or} \quad \text{column vector:} \quad \mathbf{v} = \begin{bmatrix} v_1 \\ \vdots \\ v_n \end{bmatrix}$$

We always write a vector in $\mathbb{R}^n$ as a column vector. Thus

$$\mathbb{R}^n := \left\{ \begin{bmatrix} v_1 \\ \vdots \\ v_n \end{bmatrix} : v_1, \dots, v_n \in \mathbb{R} \right\}.$$

Transpose: $[v_1, \dots, v_n]^\top = \begin{bmatrix} v_1 \\ \vdots \\ v_n \end{bmatrix}$ and $\begin{bmatrix} v_1 \\ \vdots \\ v_n \end{bmatrix}^\top = [v_1, \dots, v_n]$.

Vectors in $\mathbb{C}^n$

We define $\mathbb{C}^n$ to be the set of all ordered n-tuples of complex numbers. Thus an n-tuple in $\mathbb{C}^n$ (also called an n-vector) is of the form

$$\text{row vector:} \quad \mathbf{v} = [v_1, \ldots, v_n] \;\text{ or }\; \text{column vector:} \quad \mathbf{v} = \begin{bmatrix} v_1 \\ \vdots \\ v_n \end{bmatrix}$$

We always write a vector in $\mathbb{C}^n$ as column vector. Thus

$$\mathbb{C}^n := \left\{ \begin{bmatrix} v_1 \\ \vdots \\ v_n \end{bmatrix} : v_1, \ldots, v_n \in \mathbb{C} \right\}$$

Conjugate transpose: Here $\bar{z}$ is the complex conjugate of $z \in \mathbb{C}$.

$$[v_1, \ldots, v_n]^* = \begin{bmatrix} \bar{v}_1 \\ \vdots \\ \bar{v}_n \end{bmatrix} \text{ and } \begin{bmatrix} v_1 \\ \vdots \\ v_n \end{bmatrix}^* = [\bar{v}_1, \ldots, \bar{v}_n].$$

Algebraic properties of vectors in $\mathbb{R}^n$ and $\mathbb{C}^n$

Define addition and scalar multiplication on $\mathbb{F}^n$ ($\mathbb{F} = \mathbb{R}$ or $\mathbb{C}$) as follows:

$$\begin{bmatrix} u_1 \\ \vdots \\ u_n \end{bmatrix} + \begin{bmatrix} v_1 \\ \vdots \\ v_n \end{bmatrix} = \begin{bmatrix} u_1 + v_1 \\ \vdots \\ u_n + v_n \end{bmatrix} \text{ and } \alpha \begin{bmatrix} u_1 \\ \vdots \\ u_n \end{bmatrix} = \begin{bmatrix} \alpha u_1 \\ \vdots \\ \alpha u_n \end{bmatrix} \text{ for } \alpha \in \mathbb{F}.$$

This produces new vectors from old vectors. For $\mathbf{u}, \mathbf{v}, \mathbf{w}$ in $\mathbb{F}^n$ and scalars α, β in $\mathbb{F}$, the following hold:

1. Commutativity: $\mathbf{u} + \mathbf{v} = \mathbf{v} + \mathbf{u}$
2. Associativity: $(\mathbf{u} + \mathbf{v}) + \mathbf{w} = \mathbf{u} + (\mathbf{v} + \mathbf{w})$
3. Identity: $\mathbf{u} + \mathbf{0} = \mathbf{u}$
4. Inverse: $\mathbf{u} + (-\mathbf{u}) = \mathbf{0}$
5. Distributivity : $\alpha(\mathbf{u} + \mathbf{v}) = \alpha\mathbf{u} + \alpha\mathbf{v}$
6. Distributivity : $(\alpha + \beta)\mathbf{u} = \alpha\mathbf{u} + \beta\mathbf{u}$
7. Associativity: $\alpha(\beta\mathbf{u}) = (\alpha\beta)\mathbf{u}$
8. Identity: $1\mathbf{u} = \mathbf{u}$.

Examples of vectors

Standard vectors: The vectors
$\mathbf{e}_1 := \begin{bmatrix} 1 & 0 & \cdots 0 \end{bmatrix}^\top$, $\mathbf{e}_2 := \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 \end{bmatrix}^\top, \ldots, \mathbf{e}_n := \begin{bmatrix} 0 & \cdots & 0 & 1 \end{bmatrix}^\top$ are called standard vectors or canonical vectors in $\mathbb{R}^n$ and $\mathbb{C}^n$.

Features vectors. A feature vector collects together n different quantities that pertain to a single thing or object. The entries of a feature vector are called the features or attributes.

For instance, a 5-vector $\mathbf{x} := [x_1, x_2, x_3, x_4, x_5]^\top$ could give the age, height, weight, blood pressure, and temperature of a patient admitted to a hospital.

Word count vector. An n-vector $\mathbf{w}$ can represent the number of times each word in a dictionary of n words appears in a document.

For instance, the word count vector $[25, 2, 0]^\top$ means that the first dictionary word appears 25 times, the second one twice, and the third one not at all.

Matrices

Definition: A matrix is an array of numbers. An $m \times n$ matrix A has m rows and n columns and is of the form

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

The j-th column of A : $\mathbf{a}_j := \begin{bmatrix} a_{1j} \\ \vdots \\ a_{mj} \end{bmatrix}$ for $j = 1 : n$.

The i-th row of A: $\hat{\mathbf{a}}_i := \begin{bmatrix} a_{i1} & a_{i2} & \cdots & a_{in} \end{bmatrix}$ for $i = 1 : m$. Then

$$A = \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{bmatrix} = \left[\ \mathbf{a}_1 \mid \mathbf{a}_2 \mid \cdots \mid \mathbf{a}_n\ \right] = \begin{bmatrix} - \hat{\mathbf{a}}_1 - \\ \vdots \\ - \hat{\mathbf{a}}_m - \end{bmatrix}.$$

Special matrices

An $m \times n$ matrix said to be a square matrix if $m = n$. An $m \times n$ matrix $D := [d_{ij}]$ is said to be a diagonal matrix if $d_{ij} = 0$ for all $i \neq j$. An $n \times n$ diagonal matrix D with diagonal entries $d_1, \ldots, d_n$ is given by

$$D = \mathrm{diag}(d_1, \ldots, d_n) = \begin{bmatrix} d_1 & & \\ & \ddots & \\ & & d_n \end{bmatrix}.$$

Identity matrix: An $n \times n$ diagonal matrix with all diagonal entries equal to 1 is called the identity matrix and is denoted by I_n or I.

Zero matrix: An $m \times n$ matrix with all entries 0 is called the zero matrix and is denoted by $\mathbf{O}_{m \times n}$ or simply by $\mathbf{O}$.

Example: $I := \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$ and $\mathbf{O} := \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$.

Matrix addition and scalar multiplication

Let $\mathbb{F}^{m\times n}$ denote the set of all $m \times n$ matrices with entries in $\mathbb{F}$ where $\mathbb{F} = \mathbb{R}$ or $\mathbb{F} = \mathbb{C}$. Let $A := [a_{ij}]$ and $B := [b_{ij}]$ be matrices $\in \mathbb{F}^{m\times n}$ and $\alpha \in \mathbb{F}$.

1. Matrix addition: $A + B := [a_{ij} + b_{ij}] \in \mathbb{F}^{m\times n}$.
2. Multiplication by a scalar: $\alpha A := [\alpha\, a_{ij}] \in \mathbb{F}^{m\times n}$.

Let $A := \begin{bmatrix} 1 & 4 & 0 \\ -2 & 6 & 5 \end{bmatrix}$ and $B := \begin{bmatrix} -3 & 1 & -1 \\ 0 & 0 & 2 \end{bmatrix}$. Then

$$\begin{aligned} A + B &= \begin{bmatrix} 1 & 4 & 0 \\ -2 & 6 & 5 \end{bmatrix} + \begin{bmatrix} -3 & 1 & -1 \\ 0 & 0 & 2 \end{bmatrix} = \begin{bmatrix} -2 & 5 & -1 \\ -2 & 6 & 7 \end{bmatrix} \\ 2A &= \begin{bmatrix} 2 & 8 & 0 \\ -4 & 12 & 10 \end{bmatrix} \text{ and } (-1)A = \begin{bmatrix} -1 & -4 & 0 \\ 2 & -6 & -5 \end{bmatrix}. \end{aligned}$$

Transpose and Conjugate transpose

Transpose: The transpose of an $m \times n$ matrix $A = [a_{ij}]_{m\times n}$ is the $n \times m$ matrix denoted by $A^\top$ and is given by $A^\top = [a_{ji}]_{n\times m}$.

Example: $\begin{bmatrix} 1 & 2 \\ 4 & 5 \\ 7 & 8 \end{bmatrix}^\top = \begin{bmatrix} 1 & 4 & 7 \\ 2 & 5 & 8 \end{bmatrix}$ and $\begin{bmatrix} 1+i & 2 \\ 3 & 4+5i \end{bmatrix}^\top = \begin{bmatrix} 1+i & 3 \\ 2 & 4+5i \end{bmatrix}$

Conjugate transpose: The conjugate transpose of an $m \times n$ complex matrix $A = [a_{ij}]_{m\times n}$ is the $n \times m$ matrix denoted by A^* and is given by

$$A^* = [\bar{a}_{ji}]_{n\times m} = ([\bar{a}_{ij}]_{m\times n})^\top = (\bar{A})^\top,$$

where $\bar{a}_{ij}$ is the complex conjugate of a_{ij}.

Example: $\begin{bmatrix} i & 4 & 1+i \\ 3 & 4+5i & 0 \end{bmatrix}^* = \begin{bmatrix} -i & 3 \\ 4 & 4-5i \\ 1-i & 0 \end{bmatrix}$

Transpose and conjugate transpose

Exercise: Let $A, B \in \mathbb{F}^{m\times n}$ and $\alpha \in \mathbb{F}$. Then show that

$$(a)\,(A+B)^\top = A^\top + B^\top \quad (b)\,(\alpha A)^\top = \alpha A^\top \text{ and } (\alpha A)^* = \bar{\alpha} A^* \quad (c)\,(A^\top)^\top = A.$$

Definition: Let A be an $n \times n$ matrix. Then A is said to be

1. symmetric if $A^\top = A$
2. skew-symmetric if $A^\top = -A$
3. Hermitian if $A^* = A$
4. skew-Hermitian if $A^* = -A$.

Remark: Let $A := [a_{ij}]_{n\times n}$. If $A^\top = -A$ then $a_{jj} = 0$ for $j = 1 : n$. On the other hand, if $A^* = -A$ then $\mathrm{Re}(a_{jj}) = 0$ for $j = 1 : n$.

Matrix-vector multiplication

Let $A := \begin{bmatrix} \mathbf{a}_1 & \cdots & \mathbf{a}_n \end{bmatrix} \in \mathbb{F}^{m \times n}$ and $\mathbf{x} := [x_1, \ldots, x_n]^\top \in \mathbb{F}^n$. We define the matrix-vector multiplication $A\mathbf{x}$ to be the linear combination of columns of A.

Definition: Matrix-vector multiplication

$$A\mathbf{x} = \begin{bmatrix} \mathbf{a}_1 & \cdots & \mathbf{a}_n \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = x_1\mathbf{a}_1 + \cdots + x_n\mathbf{a}_n.$$

Example:

$$\begin{bmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ 0 & -1 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = x_1 \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix} + x_2 \begin{bmatrix} 0 \\ 1 \\ -1 \end{bmatrix} + x_3 \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} x_1 \\ -x_1 + x_2 \\ -x_2 + x_3 \end{bmatrix}.$$

Matrix-vector multiplication

A row vector $\begin{bmatrix} a_{i1} & \cdots & a_{in} \end{bmatrix}$ is a $1 \times n$ matrix. Therefore

$$\begin{bmatrix} a_{i1} & \cdots & a_{in} \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = a_{i1}x_1 + \cdots + a_{in}x_n.$$

Example: Matrix-vector multiplication in two ways

$$\begin{aligned} \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 0 & 1 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} &= x_1 \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} + x_2 \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix} + x_3 \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \\ &= \begin{bmatrix} x_1 \\ x_1 + x_2 \\ x_2 + x_3 \end{bmatrix} = \begin{bmatrix} \begin{bmatrix} 1 & 0 & 0 \end{bmatrix} \mathbf{x} \\ \begin{bmatrix} 1 & 1 & 0 \end{bmatrix} \mathbf{x} \\ \begin{bmatrix} 0 & 1 & 1 \end{bmatrix} \mathbf{x} \end{bmatrix} \end{aligned}$$

Row and column oriented matrix-vector multiplication

$$\begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ldots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = x_1 \begin{bmatrix} a_{11} \\ \vdots \\ a_{m1} \end{bmatrix} + \cdots + x_n \begin{bmatrix} a_{1n} \\ \vdots \\ a_{mn} \end{bmatrix}$$

$$= \begin{bmatrix} a_{11}x_1 + \cdots + a_{1n}x_n \\ \vdots \\ a_{m1}x_1 + \cdots + a_{mn}x_n \end{bmatrix} = \begin{bmatrix} \begin{bmatrix} a_{11} & \cdots & a_{1n} \end{bmatrix} \mathbf{x} \\ \vdots \\ \begin{bmatrix} a_{m1} & \cdots & a_{mn} \end{bmatrix} \mathbf{x} \end{bmatrix}.$$

Writing $A := \begin{bmatrix} \mathbf{a}_1 & | & \cdots & | & \mathbf{a}_n \end{bmatrix}$ and $A = \begin{bmatrix} - \hat{\mathbf{a}}_1 - \\ \vdots \\ - \hat{\mathbf{a}}_m - \end{bmatrix}$, we have

$$A\mathbf{x} = x_1\mathbf{a}_1 + \cdots + x_n\mathbf{a}_n = \begin{bmatrix} a_{11}x_1 + \cdots + a_{1n}x_n \\ \vdots \\ a_{m1}x_1 + \cdots + a_{mn}x_n \end{bmatrix} = \begin{bmatrix} \hat{\mathbf{a}}_1\mathbf{x} \\ \vdots \\ \hat{\mathbf{a}}_m\mathbf{x} \end{bmatrix}.$$

Matrix-matrix multiplication

Fact: Let $A \in \mathbb{F}^{m\times n}$. Let $\mathbf{e}_i \in \mathbb{F}^m$ and $\mathbf{e}_j \in \mathbb{F}^n$ be standard unit vectors. Then

- $A\mathbf{e}_j$ is the j-th column of A.
- $\mathbf{e}_i^\top A$ is the i-th row of A.

Let $A \in \mathbb{F}^{m\times n}$ and $B := \begin{bmatrix} \mathbf{b}_1 & \cdots & \mathbf{b}_p \end{bmatrix} \in \mathbb{F}^{n\times p}$.

Definition: Define the matrix-matrix multiplication AB by

$$AB := \begin{bmatrix} A\mathbf{b}_1 & \cdots & A\mathbf{b}_p \end{bmatrix}.$$

Reason: Define AB to be the $m \times p$ matrix such that $(AB)\mathbf{x} = A(B\mathbf{x})$ for all $\mathbf{x} \in \mathbb{F}^p$. Let $C := AB$ be given by $C = \begin{bmatrix}\mathbf{c}_1 & \cdots & \mathbf{c}_p\end{bmatrix}$. Let $\mathbf{e}_j \in \mathbb{F}^p$ be the standard unit vector. Then for $j = 1 : p$, we have $B\mathbf{e}_j = \mathbf{b}_j$ and

$$\mathbf{c}_j = C\mathbf{e}_j = (AB)\mathbf{e}_j = A(B\mathbf{e}_j) = A\mathbf{b}_j \Longrightarrow C = \begin{bmatrix}A\mathbf{b}_1 & \cdots & A\mathbf{b}_p\end{bmatrix}.$$

Matrix-matrix multiplication

Let $A = \begin{bmatrix} -\hat{\mathbf{a}}_1- \\ \vdots \\ -\hat{\mathbf{a}}_m- \end{bmatrix} \in \mathbb{F}^{m\times n}$, $B := \begin{bmatrix} \mathbf{b}_1 & \cdots & \mathbf{b}_p \end{bmatrix} \in \mathbb{F}^{n\times p}$. Then

$$AB = \begin{bmatrix} A\mathbf{b}_1 & \cdots & A\mathbf{b}_p \end{bmatrix} = \begin{bmatrix} \hat{\mathbf{a}}_1\mathbf{b}_1 & \cdots & \hat{\mathbf{a}}_1\mathbf{b}_p \\ \vdots & \cdots & \vdots \\ \hat{\mathbf{a}}_m\mathbf{b}_1 & \cdots & \hat{\mathbf{a}}_m\mathbf{b}_p \end{bmatrix} = \begin{bmatrix} \hat{\mathbf{a}}_1 B \\ \vdots \\ \hat{\mathbf{a}}_m B \end{bmatrix}.$$

Thus if $A := [a_{ij}]_{m\times n}$, $B := [b_{ij}]_{n\times p}$ and $C := AB = [c_{ij}]_{m\times p}$ then

$$c_{ij} = \hat{\mathbf{a}}_i\mathbf{b}_j = \begin{bmatrix} a_{i1} & \cdots & a_{in} \end{bmatrix} \begin{bmatrix} b_{1j} \\ \vdots \\ b_{nj} \end{bmatrix} = \sum_{k=1}^{n} a_{ik} b_{kj}.$$

Remark: If A and B are $n \times n$ matrices then in general $AB \neq BA$.

Example

Let $A = \begin{bmatrix} 1 & 3 & 2 \\ 0 & -1 & 1 \end{bmatrix}$ and $B := \begin{bmatrix} 4 & -1 \\ 1 & 2 \\ 3 & 0 \end{bmatrix}$. Then

$A\mathbf{b}_1 = \begin{bmatrix} 1 & 3 & 2 \\ 0 & -1 & 1 \end{bmatrix} \begin{bmatrix} 4 \\ 1 \\ 3 \end{bmatrix} = \begin{bmatrix} 13 \\ 2 \end{bmatrix}$ and $A\mathbf{b}_2 = \begin{bmatrix} 1 & 3 & 2 \\ 0 & -1 & 1 \end{bmatrix} \begin{bmatrix} -1 \\ 2 \\ 0 \end{bmatrix} = \begin{bmatrix} 5 \\ -2 \end{bmatrix}$.

Therefore $AB = \begin{bmatrix} A\mathbf{b}_1 & A\mathbf{b}_2 \end{bmatrix} = \begin{bmatrix} 13 & 5 \\ 2 & -2 \end{bmatrix}$. On the other hand

$\widehat{\mathbf{a}}_1 B = \begin{bmatrix} 1 & 3 & 2 \end{bmatrix} \begin{bmatrix} 4 & -1 \\ 1 & 2 \\ 3 & 0 \end{bmatrix} = \begin{bmatrix} 13 & 5 \end{bmatrix}$ and $\widehat{\mathbf{a}}_2 B = \begin{bmatrix} 0 & -1 & 1 \end{bmatrix} \begin{bmatrix} 4 & -1 \\ 1 & 2 \\ 3 & 0 \end{bmatrix} = \begin{bmatrix} 2 & -2 \end{bmatrix}$.

Therefore $AB = \begin{bmatrix} \widehat{\mathbf{a}}_1 B \\ \widehat{\mathbf{a}}_2 B \end{bmatrix} = \begin{bmatrix} 13 & 5 \\ 2 & -2 \end{bmatrix} = \begin{bmatrix} A\mathbf{b}_1 & A\mathbf{b}_2 \end{bmatrix}$.

Properties of matrix multiplication

Thoerm: Let A, B and C be matrices (whose sizes are such that the indicated operations can be performed) and let α be a scalar. Then

1. Associative Law: $(AB)C = A(BC)$
2. Left Distributive Law: $A(B + C) = AB + AC$
3. Right Distributive Law: $(A + B)C = AC + BC$
4. Scalar multiplication: $\alpha(AB) = (\alpha A)B = A(\alpha B)$
5. Multiplicative identity: If A is an $m \times n$ matrix then $I_m A = A = AI_n$.

Block matrices

Definition: An $m \times n$ block matrix (or a partitioned matrix) is a matrix of the form

$$A := \begin{bmatrix} A_{11} & \cdots & A_{1n} \\ \vdots & \cdots & \vdots \\ A_{m1} & \cdots & A_{mn} \end{bmatrix}$$

where each A_{ij} is a $p_i \times q_j$ matrix for $i = 1 : m$ and $j = 1 : n$.

Then $\begin{bmatrix} A_{i1} & \cdots & A_{in} \end{bmatrix}$ is the i-th block row of A and $\begin{bmatrix} A_{1j} \\ \vdots \\ A_{mj} \end{bmatrix}$ is the j-th block column of A.

Example: $\left[\begin{array}{cc|ccc|c} 1 & 2 & 2 & 0 & 1 & 4 \\ 3 & 4 & 1 & 2 & 3 & 5 \\ \hline 5 & 7 & 2 & 7 & 8 & 8 \\ 3 & 4 & 1 & 9 & 2 & 2 \end{array}\right]$ has 2 block rows and 3 block columns.

Block matrix operations

Block matrix addition: Let $A := [A_{ij}]_{m \times n}$ and $B := [B_{ij}]_{m \times n}$ be block matrices such that size of A_{ij} = size of B_{ij} for $i = 1 : m$ and $j = 1 : n$. Then $A + B := [A_{ij} + B_{ij}]_{m \times n}$.

Block matrix multiplication: Let $A := [A_{ij}]_{m \times n}$ and $B := [B_{ij}]_{n \times p}$ be block matrices. If the matrix multiplication $C_{ij} := \sum_{k=1}^{n} A_{ik} B_{kj}$ is well defined for $i = 1 : m$ and $j = 1 : p$ then AB is an $m \times p$ block matrix given by $AB = [C_{ij}]_{m \times p}$.

Conformal partition: If an operation on block matrices A and B are well defined then A and B are said to be partitioned conformably.

Example: $\begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \begin{bmatrix} B_{11} & B_{12} & B_{13} \\ B_{21} & B_{22} & B_{23} \end{bmatrix} =$

$$\begin{bmatrix} A_{11}B_{11} + A_{12}B_{21} & A_{11}B_{12} + A_{12}B_{22} & A_{11}B_{13} + A_{12}B_{23} \\ A_{21}B_{11} + A_{22}B_{21} & A_{21}B_{12} + A_{22}B_{22} & A_{21}B_{13} + A_{22}B_{23} \end{bmatrix}.$$

Block matrix multiplication

Example:

$$\left[\begin{array}{cc|cc} 1 & 1 & 1 & 1 \\ \hline 2 & 2 & 1 & 1 \\ 3 & 3 & 2 & 2 \end{array}\right] \left[\begin{array}{cc|cc} 1 & 1 & 1 & 1 \\ 1 & 2 & 1 & 1 \\ \hline 3 & 1 & 1 & 1 \\ 3 & 2 & 1 & 2 \end{array}\right] = \left[\begin{array}{cc|cc} 8 & 6 & 4 & 5 \\ \hline 10 & 9 & 6 & 7 \\ 18 & 15 & 10 & 12 \end{array}\right]$$

$$\left[\begin{array}{cc|cc} 1 & 1 & 1 & 1 \\ 2 & 2 & 1 & 1 \\ \hline 3 & 3 & 2 & 2 \end{array}\right] \left[\begin{array}{cc|cc} 1 & 1 & 1 & 1 \\ 1 & 2 & 1 & 1 \\ \hline 3 & 1 & 1 & 1 \\ 3 & 2 & 1 & 2 \end{array}\right] = \left[\begin{array}{cc|cc} 8 & 6 & 4 & 5 \\ 10 & 9 & 6 & 7 \\ \hline 18 & 15 & 10 & 12 \end{array}\right]$$

Outer product

Given two vectors $\mathbf{x}$ and $\mathbf{y}$ in $\mathbb{R}^n$, the standard inner product of $\mathbf{x}$ and $\mathbf{y}$ is given by

$$\langle \mathbf{x}, \mathbf{y} \rangle = x_1 y_1 + \cdots + x_n y_n = \mathbf{y}^\top \mathbf{x}.$$

Outer product: The matrix product $\mathbf{x}\mathbf{y}^\top$ is an $n \times n$ matrix and is given by

$$\mathbf{x}\mathbf{y}^\top = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \begin{bmatrix} y_1 & y_2 & \cdots & y_n \end{bmatrix} = \begin{bmatrix} x_1 y_1 & x_1 y_2 & \cdots & x_1 y_n \\ x_2 y_1 & x_2 y_2 & \cdots & x_2 y_n \\ \vdots & \vdots & \ddots & \vdots \\ x_n y_1 & x_n y_2 & \cdots & x_n y_n \end{bmatrix}.$$

The product $\mathbf{x}\mathbf{y}^\top$ is called the outer product of $\mathbf{x} \in \mathbb{R}^n$ and $\mathbf{y} \in \mathbb{R}^n$.

Outer product

Example: If $\mathbf{x} := \begin{bmatrix} 4 & 1 & 3 \end{bmatrix}^\top$ and $\mathbf{y} := \begin{bmatrix} 3 & 5 & 2 \end{bmatrix}^\top$ then

$$\mathbf{x}\mathbf{y}^\top = \begin{bmatrix} 4 \\ 1 \\ 3 \end{bmatrix} \begin{bmatrix} 3 & 5 & 2 \end{bmatrix} = \begin{bmatrix} 12 & 20 & 8 \\ 3 & 5 & 2 \\ 9 & 15 & 6 \end{bmatrix}.$$

Outer product of matrices:

Let $X := \begin{bmatrix} \mathbf{x}_1 & \mathbf{x}_2 & \cdots & \mathbf{x}_n \end{bmatrix} \in \mathbb{R}^{m \times n}$ and $Y := \begin{bmatrix} \mathbf{y}_1 & \mathbf{y}_2 & \cdots & \mathbf{y}_n \end{bmatrix} \in \mathbb{R}^{p \times n}$. Then $XY^\top \in \mathbb{R}^{m \times p}$ can be written as sum of outer products of vectors

$$XY^\top = \begin{bmatrix} \mathbf{x}_1 & \mathbf{x}_2 & \cdots & \mathbf{x}_n \end{bmatrix} \begin{bmatrix} \mathbf{y}_1^\top \\ \mathbf{y}_2^\top \\ \vdots \\ \mathbf{y}_n^\top \end{bmatrix} = \mathbf{x}_1\mathbf{y}_1^\top + \mathbf{x}_2\mathbf{y}_2^\top + \cdots + \mathbf{x}_n\mathbf{y}_n^\top.$$

Floating-Point Operation (FLOP) count

Vector-vector operations: Let $\alpha \in \mathbb{R}$. Let $\mathbf{x} := \begin{bmatrix} x_1 & \cdots & x_n \end{bmatrix}^\top \in \mathbb{R}^n$ and $\mathbf{y} := \begin{bmatrix} y_1 & \cdots & y_n \end{bmatrix}^\top \in \mathbb{R}^n$. We ignore the lower order terms for flop count.

- $\mathbf{z} \leftarrow \mathbf{x} + \mathbf{y}$ and $\mathbf{d} \leftarrow \alpha \cdot \mathbf{x}$ require n flops
- $\mathbf{z} \leftarrow \alpha \cdot \mathbf{x} + \mathbf{y}$ and $s \leftarrow \langle \mathbf{x}, \mathbf{y} \rangle = \sum_{j=1}^n x_j y_j$ require $2n$ flops

Matrix-vector operations: Let $A := \begin{bmatrix} \mathbf{a}_1 & \cdots & \mathbf{a}_n \end{bmatrix} \in \mathbb{R}^{n \times n}$ and $\beta \in \mathbb{R}$.

- $\mathbf{z} \leftarrow A\mathbf{x} = x_1\mathbf{a}_1 + \cdots + x_n\mathbf{a}_n$ and $\mathbf{d} \leftarrow \alpha \cdot A\mathbf{x} + \beta \cdot \mathbf{y}$ require $2n^2$ flops
- $\mathbf{z} \leftarrow A^\top\mathbf{x} = \begin{bmatrix} \mathbf{a}_1^\top\mathbf{x} & \cdots & \mathbf{a}_n^\top\mathbf{x} \end{bmatrix}^\top$ and $\mathbf{d} \leftarrow \alpha \cdot A^\top\mathbf{x} + \beta \cdot \mathbf{y}$ require $2n^2$ flops

Matrix-matrix operations: Let $B := \begin{bmatrix} \mathbf{b}_1 & \cdots & \mathbf{b}_n \end{bmatrix} \in \mathbb{R}^{n \times n}$ and $C \in \mathbb{R}^{n \times n}$.

- $D \leftarrow AB = \begin{bmatrix} A\mathbf{b}_1 & \cdots & A\mathbf{b}_n \end{bmatrix}$ and $D \leftarrow \alpha \cdot AB + \beta \cdot C$ require $2n^3$ flops
- $D \leftarrow A^\top B$ or $D \leftarrow AB^\top$ and $D \leftarrow \alpha \cdot A^\top B + \beta \cdot C$ require $2n^3$ flops