Linear Algebra I

Vectors and Matrices as Linear Operations on Vectors

Uwe Naumann

Informatik 12:
Software and Tools for Computational Engineering

RWTH Aachen University
Outline

Objective and Learning Outcomes

Matrix Products

Vectors
   Inner / Outer Product
   Norms
   Law of Cosines
   Geometric Projection

Matrices as Linear Operators on Vector
   Symmetry
   Inverse and Orthogonality
   Vector-Induced Matrix Norms
   Singular / Eigenvalues and -vectors

Summary and Next Steps
Objective and Learning Outcomes

Matrix Products

Vectors
  Inner / Outer Product
  Norms
  Law of Cosines
  Geometric Projection

Matrices as Linear Operators on Vector
  Symmetry
  Inverse and Orthogonality
  Vector-Induced Matrix Norms
  Singular / Eigenvalues and -vectors

Summary and Next Steps
Objective

▶ Introduction of essential concepts and terminology

Learning Outcomes

▶ You will understand
  ▶ matrix products
  ▶ vector norms and projections
  ▶ matrices as linear operators on vectors.

▶ You will be able to
  ▶ visualize concepts in \( \mathbb{R}^2 \)
  ▶ apply them to further topics in linear algebra.
Outline

Objective and Learning Outcomes

Matrix Products

Vectors
  Inner / Outer Product
  Norms
  Law of Cosines
  Geometric Projection

Matrices as Linear Operators on Vector
  Symmetry
  Inverse and Orthogonality
  Vector-Induced Matrix Norms
  Singular / Eigenvalues and -vectors

Summary and Next Steps
Matrix Products

Matrix $\times$ Vector / Matrix $\times$ Matrix

For $A = (a_{i,j}) \in \mathbb{R}^{m \times n}$, $v = (v_j) \in \mathbb{R}^n$ we define the matrix $\times$ vector product

$$y = A \cdot v = (y_i)_{i=0,\ldots,m-1} = \left( \sum_{j=0,\ldots,n-1} a_{i,j} \cdot v_j \right)_{i=0,\ldots,m-1} \in \mathbb{R}^m$$

For $A = (a_{i,j}) \in \mathbb{R}^{m \times n}$, $B = (b_{j,k}) \in \mathbb{R}^{n \times p}$ we define the matrix $\times$ matrix product

$$C = A \cdot B = (c_{i,k})_{i=0,\ldots,m-1, k=0,\ldots,p-1} = \left( \sum_{j=0,\ldots,n-1} a_{i,j} \cdot b_{j,k} \right)_{i=0,\ldots,m-1, k=0,\ldots,p-1} \in \mathbb{R}^{m \times p}$$
Outline

Objective and Learning Outcomes

Matrix Products

Vectors
- Inner / Outer Product
- Norms
- Law of Cosines
- Geometric Projection

Matrices as Linear Operators on Vector
- Symmetry
- Inverse and Orthogonality
- Vector-Induced Matrix Norms
- Singular / Eigenvalues and -vectors

Summary and Next Steps
Following the definition of matrix multiplication, two vector products can be defined.

For \( \mathbf{a}, \mathbf{b} \in \mathbb{R}^n \) the inner (scalar) vector product \( \mathbf{c} \in \mathbb{R} \) is defined as

\[
\mathbf{c} = \langle \mathbf{a}, \mathbf{b} \rangle = \mathbf{a}^T \cdot \mathbf{b} = \mathbf{b}^T \cdot \mathbf{a} = \langle \mathbf{b}, \mathbf{a} \rangle.
\]

For \( \mathbf{a} \in \mathbb{R}^m, \mathbf{b} \in \mathbb{R}^n \) the outer (dyadic) vector product \( \mathbf{C} \in \mathbb{R}^{m \times n} \) is defined as

\[
\mathbf{C} = \mathbf{a} \cdot \mathbf{b}^T = (\mathbf{b}^T \cdot \mathbf{a})^T.
\]

\( \mathbf{C} \) has (both column and row) rank 1.
The magnitude of a vector \( \mathbf{v} \in \mathbb{R}^m \) is measured by its norms defined as

\[
\| \mathbf{v} \|_k = \left( \sum_{i=0}^{m-1} |v_i|^k \right)^{\frac{1}{k}},
\]
e.g.,

1-norm: \( \| \mathbf{v} \|_1 = \sum_{i=0}^{m-1} |v_i| \)

2-norm: \( \| \mathbf{v} \|_2 = \sqrt{\sum_{i=0}^{m-1} |v_i|^2} = \sqrt{\sum_{i=0}^{m} v_i^2} = \sqrt{\mathbf{v}^T \cdot \mathbf{v}} = \sqrt{\langle \mathbf{v}, \mathbf{v} \rangle} \)

\( \infty \) - norm: \( \| \mathbf{v} \|_\infty = \max_{0 \leq i \leq m-1} |v_i| \)
Vector Norms

Properties

All vector norms \( \| \cdot \| : \mathbb{R}^m \to \mathbb{R} \)

1. are positive, i.e,

\[ \forall \mathbf{v} \in \mathbb{R}^m : \| \mathbf{v} \| \geq 0; \quad \| \mathbf{v} \| = 0 \iff \mathbf{v} = \mathbf{0}_m \]

2. are homogeneous, i.e,

\[ \forall \alpha \in \mathbb{R} : \| \alpha \cdot \mathbf{v} \| = |\alpha| \cdot \| \mathbf{v} \| \]

3. satisfy the triangle inequality (are subadditive), i.e,

\[ \forall \mathbf{u}, \mathbf{v} \in \mathbb{R}^m : \| \mathbf{u} + \mathbf{v} \| \leq \| \mathbf{u} \| + \| \mathbf{v} \| . \]

We assume \( \| \mathbf{v} \| = \| \mathbf{v} \|_2 \) (length of \( \mathbf{v} \)) unless stated otherwise.
Operations on vectors are characterized by their effect on norms and relative positions, i.e. the angle spanned by two vectors in $\mathbb{R}^n$. The following law of cosines turns out to be fundamental.

$$\|a - b\|^2 = \|a\|^2 + \|b\|^2 - 2 \cdot \|a\| \cdot \|b\| \cdot \cos(\theta)$$

For the picture on the right, Pythagoras yields $\|c\|^2 = (\|b\| - d)^2 + h^2$ and $\|a\|^2 = d^2 + h^2$ implying

$$\|c\|^2 = \|a\|^2 + \|b\|^2 - 2 \cdot \|b\| \cdot d$$

and with $d = \|a\| \cdot \cos(\theta)$ proving the law.

A similar argument holds for $h$ outside of the triangle spanned by $a$, $b$ and $c$. 
Vectors

Law of Cosines and Inner Product

The angle $\theta$ spanned by two vectors $\mathbf{a}, \mathbf{b} \in \mathbb{R}^m$ is characterized by

$$\cos(\theta) = \frac{\langle \mathbf{a}, \mathbf{b} \rangle}{\| \mathbf{a} \| \cdot \| \mathbf{b} \|}.$$

The above follows from the law of cosines with

$$\| \mathbf{a} - \mathbf{b} \|^2 = \langle \mathbf{a} - \mathbf{b}, \mathbf{a} - \mathbf{b} \rangle = \sum_{i=0}^{m-1} a_i^2 + b_i^2 - 2 \cdot a_i \cdot b_i$$

$$= \sum_{i=0}^{m-1} a_i^2 + \sum_{i=0}^{m-1} b_i^2 - 2 \cdot \sum_{i=0}^{m-1} a_i \cdot b_i = \| \mathbf{a} \|^2 + \| \mathbf{b} \|^2 - 2 \cdot \langle \mathbf{a}, \mathbf{b} \rangle$$

$$= \| \mathbf{a} \|^2 + \| \mathbf{b} \|^2 - 2 \cdot \| \mathbf{a} \| \cdot \| \mathbf{b} \| \cdot \cos(\theta)$$
Vectors

Geometric Projection

From the law of cosines follows immediately, that the scalar projection of \( a \in \mathbb{R}^m \) onto \( b \in \mathbb{R}^m \) is given by

\[
a_b = \frac{\langle a, b \rangle}{\|b\|} = \cos(\theta) \cdot \|a\|.
\]

The corresponding vector projection of is given by

\[
a_b = a_b \cdot \frac{b}{\|b\|}
\]

(vector of length \( a_b \) pointing into same direction as \( b \)).

The scalar projection of \( a \in \mathbb{R}^m \) onto the \( i \)-th Cartesian basis vector \( e_i \in \mathbb{R}^m \) is given by \( a_{e_i} = a_i \).

The corresponding vector projection is given by \( a_{e_i} = a_i \cdot e_i \).
Outline

Objective and Learning Outcomes

Matrix Products

Vectors
  Inner / Outer Product
  Norms
  Law of Cosines
  Geometric Projection

Matrices as Linear Operators on Vector
  Symmetry
  Inverse and Orthogonality
  Vector-Induced Matrix Norms
  Singular / Eigenvalues and -vectors

Summary and Next Steps
Matrices as Linear Operators on Vectors

Introduction

Matrices $A \in \mathbb{R}^{m \times n}$ induce linear functions $\dot{F} : \mathbb{R}^n \rightarrow \mathbb{R}^m$:

$$\dot{y} = \dot{F}(\dot{v}) \equiv A \cdot \dot{v}$$

and $\bar{F} : \mathbb{R}^m \rightarrow \mathbb{R}^n$:

$$\bar{v} = \bar{F}(\bar{y}) \equiv A^T \cdot \bar{y}$$

Linearity follows immediately from

$$\dot{F} (\dot{v}_1 + \dot{v}_2) = A \cdot (\dot{v}_1 + \dot{v}_2) = A \cdot \dot{v}_1 + A \cdot \dot{v}_2 = \dot{F}(\dot{v}_1) + \dot{F}(\dot{v}_2)$$

and

$$\dot{F}(\alpha \cdot \dot{v}) = A \cdot \alpha \cdot \dot{v} = \alpha \cdot A \cdot \dot{v} = \alpha \cdot \dot{F}(\dot{v})$$

(and similarly for $\bar{F}$).

Properties of matrices are defined in terms of their actions as linear operators on vectors.
A matrix $A \in \mathbb{R}^{n \times n}$ is symmetric if

$$A^T = A.$$  

It follows that

$$y = \tilde{F}(v) = A \cdot v = A^T \cdot v = \bar{F}(v).$$

For $A \in \mathbb{R}^{m \times n}$ both $A^T \cdot A$ and $A \cdot A^T$ are symmetric as

$$(A^T \cdot A)^T = A^T \cdot A^T T = A^T \cdot A$$

$$(A \cdot A^T)^T = A^T T \cdot A^T = A \cdot A^T.$$  

Both matrices play fundamental roles in data analysis.
The inverse \( A^{-1} \in \mathbb{R}^{n \times n} \) of an invertible (also: regular or non-singular) matrix \( A \in \mathbb{R}^{n \times n} \) is defined by

\[
A^{-1} \cdot A = I_n = A \cdot A^{-1},
\]

where \( I_n \in \mathbb{R}^{n \times n} \) denotes the identity in \( \mathbb{R}^n \) mapping all vectors onto themselves.

A matrix is invertible if both its rows and its columns are linearly independent, i.e, no row / column can be written as a linear combination (weighted sum) of other rows / columns.
A matrix $A \in \mathbb{R}^{n \times n}$ is orthogonal if

$$A^{-1} = A^T.$$ 

$A \cdot \mathbf{v}$ amounts to a rotation of the coordinate system; i.e. angles $\theta$ between vectors $\mathbf{v}_1, \mathbf{v}_2 \in \mathbb{R}^n$ are preserved as

$$\cos(\theta) = \frac{\langle A \cdot \mathbf{v}_1, A \cdot \mathbf{v}_2 \rangle}{\|A \cdot \mathbf{v}_1\| \cdot \|A \cdot \mathbf{v}_2\|} = \frac{\mathbf{v}_1^T \cdot A^T \cdot A \cdot \mathbf{v}_2}{\sqrt{\langle A \cdot \mathbf{v}_1, A \cdot \mathbf{v}_1 \rangle} \cdot \sqrt{\langle A \cdot \mathbf{v}_2, A \cdot \mathbf{v}_2 \rangle}}$$

$$= \frac{\mathbf{v}_1^T \cdot A^{-1} \cdot A \cdot \mathbf{v}_2}{\sqrt{\mathbf{v}_1^T \cdot A^T \cdot A \cdot \mathbf{v}_1} \cdot \sqrt{\mathbf{v}_2^T \cdot A^T \cdot A \cdot \mathbf{v}_2}}$$

$$= \frac{\mathbf{v}_1^T \cdot \mathbf{v}_2}{\sqrt{\mathbf{v}_1^T \cdot A^{-1} \cdot A \cdot \mathbf{v}_1} \cdot \sqrt{\mathbf{v}_2^T \cdot A^{-1} \cdot A \cdot \mathbf{v}_2}}$$

$$= \frac{\mathbf{v}_1^T \cdot \mathbf{v}_2}{\sqrt{\mathbf{v}_1^T \cdot \mathbf{v}_1} \cdot \sqrt{\mathbf{v}_2^T \cdot \mathbf{v}_2}} = \frac{\langle \mathbf{v}_1, \mathbf{v}_2 \rangle}{\|\mathbf{v}_1\| \cdot \|\mathbf{v}_2\|}.$$
Matrices are Linear Operators on Vectors
Orthogonality

and norms are preserved as, e.g,

\[ \|A \cdot v\| = \sqrt{\langle A \cdot v, A \cdot v \rangle} = \sqrt{v^T \cdot A^T \cdot A \cdot v} \]
\[ = \sqrt{v^T \cdot A^{-1} \cdot A \cdot v} = \sqrt{v^T \cdot v} = \sqrt{\langle v, v \rangle} \]
\[ = \|v\| \]

A is called orthonormal if \(\|A\| = 1\).

If \(A\) is orthogonal and symmetric then

\[ A^{-1} = A^T = A \]
Matrices are Linear Operators on Vectors

Vector-Induced Matrix Norms

The magnitude of matrices $A \in \mathbb{R}^{m \times n}$ is typically measured in terms of norms derived from (induced by) the corresponding vector norms, e.g,

- **1-norm** (maximum absolute column sum)
  \[
  \|A\|_1 = \max_{j=0,\ldots,n-1} \sum_{i=0}^{m-1} |a_{i,j}|
  \]

- **2-norm** (maximum singular value / eigenvalue of $A^T \cdot A$)
  \[
  \|A\|_2 = \sqrt{\lambda_{\text{max}} (A^T \cdot A)}
  \]

- **∞-norm** (maximum absolute row sum)
  \[
  \|A\|_\infty = \max_{i=0,\ldots,m-1} \sum_{j=0}^{n-1} |a_{i,j}|
  \]

Vector-induced matrix norms can be shown to be submultiplicative, i.e,

\[
\forall A \in \mathbb{R}^{m \times n}, B \in \mathbb{R}^{n \times k} : \|A \cdot B\| \leq \|A\| \cdot \|B\|.
\]
Vector-Induced Matrix Norms

Matrix Norms ⇒ Vector Norms

For \( n = 1 \) the matrix \( A = (a_{i,j}) \in \mathbb{R}^{m \times 1} \) becomes a vector \( a = (a_i) \in \mathbb{R}^m \). The matrix norms turn out to be equivalent to the corresponding vector norm, e.g,

1-norm

\[
\| A \|_1 = \sum_{i=0}^{m-1} |a_{i,0}| = \sum_{i=0}^{m-1} |a_i| = \| a \|_1
\]

2-norm

\[
\| A \|_2 = \sqrt{\lambda_{\text{max}} (A^T \cdot A)} = \sqrt{\lambda_{\text{max}} (a^T \cdot a)} = \sqrt{a^T \cdot a} = \| a \|_2
\]

∞-norm

\[
\| A \|_\infty = \max_{i=0,\ldots,m-1} |a_{i,0}| = \max_{i=0,\ldots,m-1} |a_i| = \| a \|_\infty
\]
Matrices are Linear Operators on Vectors

Singular / Eigenvalues and -vectors

Relevant special cases of the impact of matrices $A \in \mathbb{R}^{n \times n}$ as linear operators acting on vectors $\mathbf{v} \in \mathbb{R}^n$ are characterized as pairs of eigenvalues $\lambda \in \mathbb{R}$ and -vectors $\mathbf{v} \in \mathbb{R}^n$ satisfying the equality

$$A \cdot \mathbf{v} = \lambda \cdot \mathbf{v}.$$  

Eigenvectors of $A$ are simply scaled by the magnitude of the associated eigenvalue. Negative eigenvalues yield reversal of the orientation of the associated eigenvectors $\mathbf{v}$.

For symmetric matrices $A \in \mathbb{R}^{n \times n}$ ($A = A^T$) one can show that all eigenvalues are non-complex ($\in \mathbb{R}$) and that the largest absolute eigenvalue $|\lambda_{\text{max}}|$ quantifies the maximum stretching of any vector $\mathbf{v}$ under $A$, i.e,

$$\forall \mathbf{v} \in \mathbb{R}^n : \|A \cdot \mathbf{v}\| \leq |\lambda_{\text{max}}| \cdot \|\mathbf{v}\|.$$  

For $B \in \mathbb{R}^{m \times n}$ eigenvalues and -vectors of $B^T \cdot B$ and $B \cdot B^T$ are referred to a singular values and vectors.
Outline

Objective and Learning Outcomes

Matrix Products

Vectors
  Inner / Outer Product
  Norms
  Law of Cosines
  Geometric Projection

Matrices as Linear Operators on Vector
  Symmetry
  Inverse and Orthogonality
  Vector-Induced Matrix Norms
  Singular / Eigenvalues and -vectors

Summary and Next Steps
Summary

▶ Essential terminology and concepts
  ▶ Vectors
  ▶ Matrices as linear operations on vectors

Next Steps

▶ Visualize concepts in $\mathbb{R}^2$.
▶ Refer to literature for further details.
▶ Continue the course to find out more...