Knowledge Discovery and Data Mining 1 (VO) (707.003)

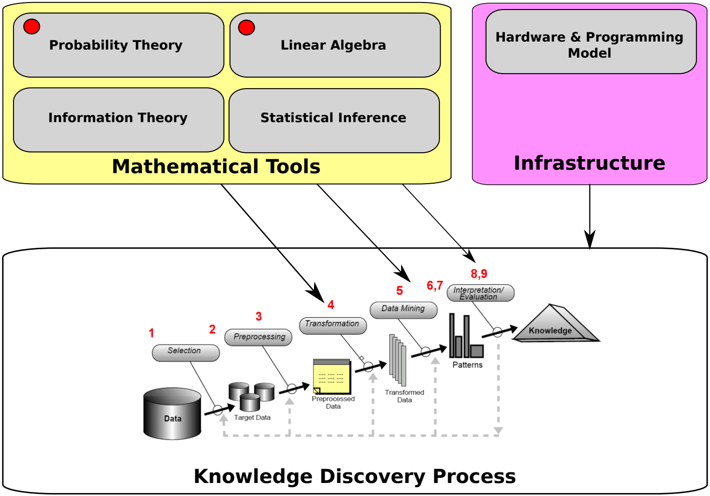
Review of Linear Algebra

Denis Helic

KTI, TU Graz

Oct 9, 2014

Big picture: KDDM



Outline

Matrix

Matrix

Matrix is a rectangular array of numbers, i.e. $\mathbf{A} \in \mathbb{R}^{m \times n}$

$$\mathbf{A} = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix}$$

Matrix: Examples

$$\mathbf{A} = \begin{pmatrix} 23 & 12 & 56 \\ 18 & 89 & 45 \end{pmatrix} \qquad \mathbf{B} = \begin{pmatrix} 45 & 17 \\ 22 & 15 \\ -18 & 14 \\ 21 & -13.6 \end{pmatrix}$$

- $\mathbf{A} \in \mathbb{R}^{2\times 3}$
- $\mathbf{B} \in \mathbb{R}^{4\times 2}$
- m and n are matrix dimensions

Multiplication with a scalar

Scalar multiplication

Given a matrix $\mathbf{A} \in \mathbb{R}^{m \times n}$ and a scalar $\alpha \in \mathbb{R}$ the scalar multiplication of $\mathbf{A}$ with α is the matrix $\mathbf{B} = \alpha\mathbf{A} \in \mathbb{R}^{m \times n}$, where

$$B_{ij} = \alpha A_{ij}$$

$$\alpha\mathbf{A} = \begin{pmatrix} \alpha a_{11} & \alpha a_{12} & \dots & \alpha a_{1n} \\ \alpha a_{21} & \alpha a_{22} & \dots & \alpha a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \alpha a_{m1} & \alpha a_{m2} & \dots & \alpha a_{mn} \end{pmatrix}$$

Multiplication with a scalar: example

$$2 \cdot \begin{pmatrix} 23 & 12 & 56 \\ 18 & 89 & 45 \end{pmatrix} = \begin{pmatrix} 46 & 24 & 112 \\ 36 & 178 & 90 \end{pmatrix}$$

Vector

Vector

Vector is a matrix consisting of one column, i.e. $\mathbf{x} \in \mathbb{R}^n$. A row vector is a matrix consisting of one row.

$$\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$$

Vector: Example

$$\mathbf{x} = \begin{pmatrix} 10 \\ 12 \\ 15 \\ -100 \end{pmatrix}$$

- $\mathbf{x} \in \mathbb{R}^4$
- n is the vector dimension

Matrix Multiplication

Matrix Multiplication

The product of two matrices $\mathbf{A} \in \mathbb{R}^{m \times n}$ and $\mathbf{B} \in \mathbb{R}^{n \times p}$ is the matrix $\mathbf{C} = \mathbf{AB} \in \mathbb{R}^{m \times p}$, where

$$C_{ij} = \sum_{k=1}^{n} A_{ik} B_{kj} \tag{1}$$

- For the product to exist: the number of columns in **A** must equal the number of rows in **B**

Matrix Multiplication: Example

$$\begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix} \begin{pmatrix} 4 & 5 \\ 10 & 2 \\ 2 & 10 \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$$

Matrix Multiplication: Example

$$\begin{pmatrix} \color{red}{1} & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix} \begin{pmatrix} \color{red}{4} & 5 \\ 10 & 2 \\ 2 & 10 \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$$

$$a_{11} = 1 \cdot 4 +$$

Matrix Multiplication: Example

$$\begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix} \begin{pmatrix} 4 & 5 \\ 10 & 2 \\ 2 & 10 \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$$

$$a_{11} = 1 \cdot 4 + 2 \cdot 10 +$$

Matrix Multiplication: Example

$$\begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix} \begin{pmatrix} 4 & 5 \\ 10 & 2 \\ 2 & 10 \end{pmatrix} = \begin{pmatrix} 30 & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$$

$$a_{11} = 1 \cdot 4 + 2 \cdot 10 + 3 \cdot 2 = 30$$

Matrix Multiplication: Example

$$\begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix} \begin{pmatrix} 4 & 5 \\ 10 & 2 \\ 2 & 10 \end{pmatrix} = \begin{pmatrix} 30 & 39 \\ a_{21} & a_{22} \end{pmatrix}$$

$$a_{12} = 1 \cdot 5 + 2 \cdot 2 + 3 \cdot 10 = 39$$

Matrix Multiplication: Example

$$\begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix} \begin{pmatrix} 4 & 5 \\ 10 & 2 \\ 2 & 10 \end{pmatrix} = \begin{pmatrix} 30 & 39 \\ 34 & a_{22} \end{pmatrix}$$

$$a_{21} = 3 \cdot 4 + 2 \cdot 10 + 1 \cdot 2 = 34$$

Matrix Multiplication: Example

$$\begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix} \begin{pmatrix} 4 & 5 \\ 10 & 2 \\ 2 & 10 \end{pmatrix} = \begin{pmatrix} 30 & 39 \\ 34 & 29 \end{pmatrix}$$

$$a_{22} = 3 \cdot 5 + 2 \cdot 2 + 1 \cdot 10 = 29$$

Vector-Vector products

Inner product

Given two vectors $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ the quantity $\mathbf{x}^T\mathbf{y}$, sometimes called the inner product or dot product, is a real number given by

$$\mathbf{x}^T\mathbf{y} = \begin{pmatrix} x_1 & x_2 & \dots x_n \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \\ \vdots y_n \end{pmatrix} = \sum_{i=1}^{n} x_i y_i$$

- The inner products are just special cases of matrix multiplication

Inner product: Example

$$\begin{pmatrix} 3 & 2 & 1 \end{pmatrix} \begin{pmatrix} 5 \\ 2 \\ 10 \end{pmatrix} = 29$$

$$\mathbf{x}^T \mathbf{y} = 3 \cdot 5 + 2 \cdot 2 + 1 \cdot 10 = 29$$

Vector-Vector products

Outer product

Given two vectors $\mathbf{x} \in \mathbb{R}^m$ and $\mathbf{y} \in \mathbb{R}^n$ the matrix $\mathbf{x}\mathbf{y}^T$ is called the outer product of the vectors. The elements of the matrix are given by $(\mathbf{x}\mathbf{y}^T)_{ij} = x_i y_j$, i.e.

$$\mathbf{x}\mathbf{y}^T = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_m \end{pmatrix} \begin{pmatrix} y_1 & y_2 & \dots & y_n \end{pmatrix} = \begin{pmatrix} x_1 y_1 & x_1 y_2 & \dots & x_1 y_n \\ x_2 y_1 & x_2 y_2 & \dots & x_2 y_n \\ \vdots & \vdots & \ddots & \vdots \\ x_m y_1 & x_m y_2 & \dots & x_m y_n \end{pmatrix}$$

Matrix-Vector products

Matrix-vector product

Given a matrix $\mathbf{A} \in \mathbb{R}^{m \times n}$ and a vector $\mathbf{x} \in \mathbb{R}^n$, their product is a vector $\mathbf{y} = \mathbf{A}\mathbf{x} \in \mathbb{R}^m$, i.e.

$$\mathbf{A}\mathbf{x} = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} \sum_{i=1}^{n} a_{1i} x_i \\ \sum_{i=1}^{n} a_{2i} x_i \\ \vdots \\ \sum_{i=1}^{n} a_{mi} x_i \end{pmatrix} = \begin{pmatrix} \mathbf{a}_1^T \mathbf{x} \\ \mathbf{a}_2^T \mathbf{x} \\ \vdots \\ \mathbf{a}_m^T \mathbf{x} \end{pmatrix}$$

Matrix-Vector products: Example

$$\begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix} \begin{pmatrix} 4 \\ 10 \\ 2 \end{pmatrix} = \begin{pmatrix} 30 \\ y_2 \end{pmatrix}$$

$$y_1 = 1 \cdot 4 + 2 \cdot 10 + 3 \cdot 2 = 30$$

Matrix-Vector products: Example

$$\begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix} \begin{pmatrix} 4 \\ 10 \\ 2 \end{pmatrix} = \begin{pmatrix} 30 \\ 34 \end{pmatrix}$$

$$y_2 = 3 \cdot 4 + 2 \cdot 10 + 1 \cdot 2 = 34$$

Matrix-Vector products

Matrix-vector product

Given a matrix $\mathbf{A} \in \mathbb{R}^{m \times n}$ and a vector $\mathbf{x} \in \mathbb{R}^m$, their (left row) product is a vector $\mathbf{y}^t = \mathbf{x}^T\mathbf{A} \in \mathbb{R}^n$, i.e.

$$
\begin{aligned}
\mathbf{x}^T\mathbf{A} &= \begin{pmatrix} x_1 & x_2 & \dots & x_m \end{pmatrix} \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix} \\
&= \begin{pmatrix} \sum_{i=1}^m a_{i1}x_i & \sum_{i=1}^m a_{i2}x_i & \dots & \sum_{i=1}^m a_{in}x_i \end{pmatrix} = \begin{pmatrix} \mathbf{x}^T\mathbf{a}_1 & \mathbf{x}^T\mathbf{a}_2 & \dots & \mathbf{x}^T\mathbf{a}_m \end{pmatrix}
\end{aligned}
$$

Matrix-Vector products: Example

$$\begin{pmatrix} 4 & 10 \end{pmatrix} \begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix} = \begin{pmatrix} 34 & y_2 & y_3 \end{pmatrix}$$

$$y_1 = 4 \cdot 1 + 10 \cdot 3 = 34$$

Matrix-Vector products: Example

$$\begin{pmatrix} 4 & 10 \end{pmatrix} \begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix} = \begin{pmatrix} 34 & 28 & y_3 \end{pmatrix}$$

$$y_2 = 4 \cdot 2 + 10 \cdot 2 = 28$$

Matrix-Vector products: Example

$$\begin{pmatrix} 4 & 10 \end{pmatrix} \begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix} = \begin{pmatrix} 34 & 28 & 22 \end{pmatrix}$$

$$y_3 = 4 \cdot 3 + 10 \cdot 1 = 22$$

Properties of matrix multiplication

- Matrix multiplication is associative: $(\mathbf{AB})\mathbf{C} = \mathbf{A}(\mathbf{BC})$
- Matrix multiplication is distributive: $\mathbf{A}(\mathbf{B} + \mathbf{C}) = \mathbf{AB} + \mathbf{AC}$
- Matrix multiplication is not commutative: $\mathbf{AB} \neq \mathbf{BA}$

Matrix Sum

Matrix Sum

The sum of two matrices $\mathbf{A} \in \mathbb{R}^{m \times n}$ and $\mathbf{B} \in \mathbb{R}^{m \times n}$ is the matrix $\mathbf{C} = \mathbf{A} + \mathbf{B} \in \mathbb{R}^{m \times n}$, where

$$C_{ij} = A_{ij} + B_{ij} \tag{2}$$

- For the sum to exist: the dimension of **A** and **B** must be equal

Matrix Sum: Example

$$\begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix} + \begin{pmatrix} 4 & 5 & 10 \\ 2 & 2 & 10 \end{pmatrix} = \begin{pmatrix} 5 & 7 & 13 \\ 5 & 4 & 11 \end{pmatrix}$$

Linear combinations of vectors

Linear combination

Given are k vectors $\mathbf{x}_i \in \mathbb{R}^n$ and k scalars $\alpha_i \in \mathbb{R}$. Then the linear combination of those vectors with those scalars is defined as:

$$\alpha_1\mathbf{x}_1 + \alpha_2\mathbf{x}_2 + \ldots \alpha_k\mathbf{x}_k = \sum_{i=1}^{k} \alpha_i\mathbf{x}_i \tag{3}$$

Linear combinations of vectors: example

$$\begin{pmatrix} 1 \\ 3 \\ 6 \end{pmatrix} = 1 \cdot \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + 3 \cdot \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} + 6 \cdot \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \tag{4}$$

- In fact, any vector in $\mathbb{R}^3$ is a linear combination of those three vectors

Identity matrix

Identity matrix

The identity matrix, denoted $\mathbf{I} \in \mathbb{R}^{n \times n}$ is a square matrix with ones on the diagonal and zeros everywhere else, that is

$$I_{ij} = \begin{cases} 1, & i = j \\ 0, & i \neq j \end{cases}$$

- The identity matrix has the property that for all $\mathbf{A} \in \mathbb{R}^{m \times n}$

$$\mathbf{AI} = \mathbf{A} = \mathbf{IA} \tag{5}$$

Identity matrix

$$\mathbf{I} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

Diagonal matrix

Diagonal matrix

A diagonal matrix, denoted $\mathbf{D} \in \mathbb{R}^{n \times n}$ is a square matrix with all non-diagonal elements zero. This is typically denoted $\mathbf{D} = \text{diag}(d_1, d_2, \ldots, d_n)$, with

$$D_{ij} = \begin{cases} d_i, \ i = j \\ 0, \ i \neq j \end{cases}$$

- The identity matrix is then $\mathbf{I} = \text{diag}(1, 1, \ldots, 1)$

Diagonal matrix

$$\mathbf{D} = \begin{pmatrix} 5 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 6 & 0 \\ 0 & 0 & 0 & -10 \end{pmatrix}$$

The Transpose

Transpose

The transpose of a matrix results from "flipping" the rows and columns. Given a matrix $\mathbf{A} \in \mathbb{R}^{m \times n}$, its transpose $\mathbf{A}^T \in \mathbb{R}^{n \times m}$, is the $n \times m$ matrix whose entries are given

$$A^T_{ij} = A_{ji}$$

- We have been using the transpose for describing row vectors
- Transpose of a column vector is a row vector

Properties of transpose

- $(\mathbf{A}^T)^T = \mathbf{A}$
- $(\mathbf{AB})^T = \mathbf{B}^T\mathbf{A}^T$
- $(\mathbf{A}+\mathbf{B})^T = \mathbf{A}^T + \mathbf{B}^T$

Transpose: Example

$$\begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix}^T = \begin{pmatrix} 1 & 3 \\ 2 & 2 \\ 3 & 1 \end{pmatrix}$$

Symmetric Matrices

Symmetric matrices

A square matrix $\mathbf{A} \in \mathbb{R}^{n \times n}$ is symmetric if $\mathbf{A} = \mathbf{A}^T$.

- For any matrix $\mathbf{A} \in \mathbb{R}^{n \times n}$, the matrix $\mathbf{A} = \mathbf{A}^T$ is symmetric
- Proof left for exercise ;)
- Symmetric matrices occur very often in practice and they have many nice properties

Symmetric Matrices: Example

$$\mathbf{A} = \begin{pmatrix} 0 & 1 & 0 & 0 & 1 \\ 1 & 0 & 1 & 1 & 0 \\ 0 & 1 & 0 & 1 & 1 \\ 0 & 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 \end{pmatrix} \quad (6)$$

Trace

Trace

The trace of a square matrix $\mathbf{A} \in \mathbb{R}^{n \times n}$, denoted by $tr(\mathbf{A})$ is the sum of diagonal elements in the matrix:

$$tr(\mathbf{A}) = \sum_{i=1}^{n} A_{ii} \tag{7}$$

Trace: example

$$tr(\mathbf{A}) = tr(\begin{pmatrix} 5 & 0 & 15 & 0 \\ 0 & 1 & 0 & 21 \\ 0 & -19 & 6 & 0 \\ 10 & 12 & 16 & -10 \end{pmatrix}) = 2$$

Norms

Norm

A norm of a vector $||\mathbf{x}||$ is informally a measure of the "length" of the vector. For example, we have the commonly used Euclidean or ℓ_2 norm,

$$||\mathbf{x}||_2 = \sqrt{\sum_{i=1}^{n} x_i^2} \tag{8}$$

- Note that: $||\mathbf{x}||_2^2 = \mathbf{x}^T\mathbf{x}$

Norms

Norm

A norm is any function $f : \mathbb{R}^n \to \mathbb{R}$ that satisfies 4 properties:

1. For all $\mathbf{x} \in \mathbb{R}^n$, $f(\mathbf{x}) \geq 0$ (non-negativity)
2. $f(\mathbf{x}) = 0$, if and only if $\mathbf{x} = \mathbf{0}$ (definiteness)
3. For all $\mathbf{x} \in \mathbb{R}^n$, $t \in \mathbb{R}$, $f(t\mathbf{x}) = |t|f(\mathbf{x})$ (homogeneity)
4. For all $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$, $f(\mathbf{x} + \mathbf{y}) \leq f(\mathbf{x}) + f(\mathbf{y})$ (triangle inequality)

- Norms are used to normalize vectors and obtain unit vectors (vectors of "length" 1)

Norms

- Other examples of norms include ℓ_1, ℓ_∞, or more generally ℓ_p, with $p \geq 1$

$$||\mathbf{x}||_1 = \sum_{i=1}^{n} |x_i|$$

$$||\mathbf{x}||_\infty = \max_i |x_i|$$

$$||\mathbf{x}||_p = \left(\sum_{i=1}^{n} |x_i|^p \right)^{1/p}$$

Norms: example

$$\mathbf{x} = \begin{pmatrix} 4 \\ -10 \\ 2 \end{pmatrix}$$

$$||\mathbf{x}||_1 = 16$$

$$||\mathbf{x}||_2 = 10.95$$

$$||\mathbf{x}||_1 = 10$$

Linear independence

Linear independence

A set of vectors $\{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_k\} \subset \mathbb{R}^n$ is said to be linearly independent if no vector can be represented as a linear combination of the remaining vectors.

Linear independence

A set of vectors $\{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_k\} \subset \mathbb{R}^n$ is said to be linearly independent if no scalars $\{\alpha_1, \alpha_2, \ldots, \alpha_k\}$ exist such that

$$\sum_{i=1}^{k} \alpha_i \mathbf{x}_i = \mathbf{0}$$

Linear independence: example

$$\mathbf{x}_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \qquad \mathbf{x}_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \qquad \mathbf{x}_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$$

Linear dependence

- Conversly, if one vector belonging to the set can be represented as a linear combination of the remaining vectors the vectors are linearly dependent

$$\mathbf{x}_1 = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} \qquad \mathbf{x}_2 = \begin{pmatrix} 4 \\ 1 \\ 5 \end{pmatrix} \qquad \mathbf{x}_3 = \begin{pmatrix} 2 \\ -3 \\ -1 \end{pmatrix}$$

- Dependent because: $\mathbf{x}_3 = -2\mathbf{x}_1 + \mathbf{x}_2$

Vector bases

Span

A set of n vectors $\{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n\} \subset \mathbb{R}^n$ is said to span $\mathbb{R}^n$, i.e. they form a basis for the space. Any vector $\mathbf{v} \in \mathbb{R}^n$ can be written as a linear combination of $\mathbf{x}_1$ through $\mathbf{x}_n$.

Vector bases

Orthogonal basis

An orthogonal basis $\{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n\} \subset \mathbb{R}^n$ satisfies

$$\mathbf{x}_i^T \mathbf{x}_j = 0, \text{ if } i \neq j$$

Orthonormal basis

An orthonormal basis $\{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n\} \subset \mathbb{R}^n$ satisfies

$$\begin{aligned} \mathbf{x}_i^T \mathbf{x}_j &= 0, \text{ if } i \neq j \\ \mathbf{x}_i^T \mathbf{x}_j &= 1, \text{ if } i = j \end{aligned}$$

Orthonormal basis: example

$$\mathbf{x}_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \qquad \mathbf{x}_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \qquad \mathbf{x}_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$$

- Orthonormal basis for $\mathbb{R}^3$

Rank

Rank

The column rank of a matrix $\mathbf{A} \in \mathbb{R}^{m \times n}$ is the size of the largest subset of columns that constitute a linearly independent set. The row rank is the largest subset of rows that constitute a linearly independent set.

- For any matrix $\mathbf{A} \in \mathbb{R}^{m \times n}$ the columns rank is equal to row rank and both quantities are referred to as the rank of $\mathbf{A}$, denoted as $rank(\mathbf{A})$

Rank: example

$$rank\left(\begin{pmatrix} 5 & 0 & 15 \\ 0 & 1 & 0 \\ -5 & 1 & -15 \end{pmatrix}\right) = 2$$

Properties of the matrix rank

- Some basic properties of the rank:
  1. For $\mathbf{A} \in \mathbb{R}^{m \times n}$, $rank(\mathbf{A}) \leq min(m, n)$. If $rank(\mathbf{A}) = min(m, n)$ then $\mathbf{A}$ is said to be full rank.
  2. For $\mathbf{A} \in \mathbb{R}^{m \times n}$, $rank(\mathbf{A}) = rank(\mathbf{A}^T)$
  3. For $\mathbf{A} \in \mathbb{R}^{m \times n}$, $\mathbf{B} \in \mathbb{R}^{n \times p}$, $rank(\mathbf{AB}) = min(rank(\mathbf{A}), rank(\mathbf{B}))$
  4. For $\mathbf{A}, \mathbf{B} \in \mathbb{R}^{m \times n}$, $rank(\mathbf{A} + \mathbf{B}) \leq rank(\mathbf{A}) + rank(\mathbf{B})$

The Inverse

Inverse

The inverse of a square matrix $\mathbf{A} \in \mathbb{R}^{n \times n}$ is denoted $\mathbf{A}^{-1}$ and is a unique matrix such that

$$\mathbf{A}^{-1}\mathbf{A} = \mathbf{I} = \mathbf{A}\mathbf{A}^{-1}$$

- Not all matrices have inverses
- If $\mathbf{A}^{-1}$ exists we say that $\mathbf{A}$ is invertible or non-singular
- If $\mathbf{A}^{-1}$ does not exist we say that $\mathbf{A}$ is non-invertible or singular

The Inverse

- For $\mathbf{A}$ to be invertible it must be full rank
- There are many alternative conditions for invertability
- The inverse is used in e.g. solving the system of linear equations: $\mathbf{A}\mathbf{x} = \mathbf{b}$
- With $\mathbf{A} \in \mathbb{R}^{n \times n}$ and $\mathbf{x}, \mathbf{b} \in \mathbb{R}^n$
- Then $\mathbf{x} = \mathbf{A}^{-1}\mathbf{b}$

Properties of inverse

- $(\mathbf{A}^{-1})^{-1} = \mathbf{A}$
- $(\mathbf{AB})^{-1} = \mathbf{B}^{-1}\mathbf{A}^{-1}$
- $(\mathbf{A}^{-1})^{T} = (\mathbf{A}^{T})^{-1}$

Inverse: Example

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}^{-1} = \frac{1}{ad - bc} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$$

$$\begin{pmatrix} 6 & 2 \\ 1 & 5 \end{pmatrix}^{-1} = \frac{1}{28} \begin{pmatrix} 5 & -2 \\ -1 & 6 \end{pmatrix}$$

Orthogonal matrices

Inverse

A square matrix $\mathbf{U} \in \mathbb{R}^{n \times n}$ is orthogonal if all its columns are orthonormal (orthogonal to each other and normalized). It follows from the definition:

$$\mathbf{U}^T\mathbf{U} = \mathbf{I} = \mathbf{U}\mathbf{U}^T$$

- In other words, the inverse of an orthogonal matrix is its transpose

Orthogonal matrices: Example

$$\begin{pmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{pmatrix}^T \begin{pmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{pmatrix} \begin{pmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{pmatrix}^T$$

The determinant

Determinant

The determinant is a real number associated with a square matrix $\mathbf{A} \in \mathbb{R}^{n \times n}$, denoted $det(\mathbf{A})$. Depending on n it can be calculated by different arithmetic expressions.

$$det(\begin{pmatrix} a & b \\ c & d \end{pmatrix}) = ad - bc$$

$$det(\begin{pmatrix} 6 & 2 \\ 1 & 5 \end{pmatrix}) = 28$$

Eigenvalues and eigenvectors

Eigenvalues and eigenvectors

Given a square matrix $\mathbf{A} \in \mathbb{R}^{n \times n}$, we say that $\lambda \in \mathbb{C}$ is an eigenvalue of $\mathbf{A}$ and $\mathbf{x} \in \mathbb{C}^n$ is the corresponding eigenvector if

$$\mathbf{Ax} = \lambda \mathbf{x}, \mathbf{x} \neq \mathbf{0}$$

- Intuitively, multiplying $\mathbf{A}$ by the vector $\mathbf{x}$ does not change the direction of $\mathbf{x}$
- The new vector points in the same direction but is scaled by factor λ

Eigenvalues and eigenvectors

- Also, for any eigenvector $\mathbf{x} \in \mathbb{C}^n$, and scalar $c \in \mathbb{C}$:

$$\mathbf{A}(c\mathbf{x}) = c\mathbf{A}\mathbf{x} = c\lambda\mathbf{x} = \lambda(c\mathbf{x})$$

- Thus, $c\mathbf{x}$ is also an eigenvector with eigenvalue λ
- For this reason, we usually normalize the eigenvector $\mathbf{x}$ to have length 1

Eigenvalues and eigenvectors

- We can rewrite the equation from above (it holds: $\mathbf{x} \neq \mathbf{0}$):

$$
\begin{aligned}
\mathbf{A}\mathbf{x} &= \lambda\mathbf{x} \\
\mathbf{A}\mathbf{x} &= \lambda\mathbf{I}\mathbf{x} \\
\lambda\mathbf{I}\mathbf{x} - \mathbf{A}\mathbf{x} &= 0 \\
(\lambda\mathbf{I} - \mathbf{A})\mathbf{x} &= 0
\end{aligned}
$$

Eigenvalues and eigenvectors

- $(\lambda \mathbf{I} - \mathbf{A})\mathbf{x} = 0, \mathbf{x} \neq \mathbf{0}$ has a non-zero solution only if the matrix $(\lambda \mathbf{I} - \mathbf{A})$ is singular, i.e.:

$$det(\lambda \mathbf{I} - \mathbf{A}) = 0$$

- This can be expanded in a very large polynomial in λ, with maximal degree n
- We then find the roots of that polynomial and obtain the eigenvalues $\lambda_1, \ldots, \lambda_n$

Eigenvalues and eigenvectors

- To find the eigenvector corresponding to the eigenvalue λ_i we simply solve the linear equation:

$$(\lambda_i \mathbf{I} - \mathbf{A})\mathbf{x} = 0$$

Eigenvalues and eigenvectors: example

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

$$\begin{aligned} det(\lambda \mathbf{I} - \mathbf{A}) &= det(\begin{pmatrix} \lambda - 2 & 4 \\ 1 & \lambda + 1 \end{pmatrix}) = (\lambda - 2)(\lambda + 1) - 4 \\ &= \lambda^2 - 2\lambda + \lambda - 2 - 4 = \lambda^2 - \lambda - 6 = (\lambda - 3)(\lambda + 2) \end{aligned}$$

- Thus, $\lambda_1 = 3$, and $\lambda_2 = -2$ are eigenvalues of $\mathbf{A}$
- We now solve $(\lambda_i \mathbf{I} - \mathbf{A})\mathbf{x} = 0$ for each eigenvalue to find the corresponding eigenvectors

Eigenvalues and eigenvectors: example

- For $\lambda_1 = 3$

$$\begin{pmatrix} 1 & 4 \\ 1 & 4 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

$$\begin{aligned} x_1 + 4x_2 &= 0 \\ x_1 + 4x_2 &= 0 \end{aligned}$$

- Thus, $x_1 = -4x_2$, and we might pick $\mathbf{x} = \begin{pmatrix} -4 \\ 1 \end{pmatrix}$
- And we normalize to: $\begin{pmatrix} -4/\sqrt{17} \\ 1/\sqrt{17} \end{pmatrix}$

Eigenvalues and eigenvectors: example

- For $\lambda_1 = -2$

$$\begin{pmatrix} -4 & 4 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

$$\begin{aligned} -4x_1 + 4x_2 &= 0 \\ x_1 - x_2 &= 0 \end{aligned}$$

- Thus, $x_1 = x_2$, and we might pick $\mathbf{x} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$
- And we normalize to: $\begin{pmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \end{pmatrix}$

Properties of eigenvalues and eigenvectors

- For $\mathbf{A} \in \mathbb{R}^{n \times n}$, eigenvalues $\lambda_1, \ldots, \lambda_n$, and associated eigenvectors $\mathbf{x}_1, \ldots, \mathbf{x}_n$ the following are properties:
  1. The trace of $\mathbf{A}$ is equal to the sum of its eigenvalues:

$$tr(\mathbf{A}) = \sum_{i=1}^{n} \lambda_i$$

  2. The determinant of $\mathbf{A}$ is equal to the product of its eigenvalues:

$$det(\mathbf{A}) = \prod_{i=1}^{n} \lambda_i$$

Properties of eigenvalues and eigenvectors

- For $\mathbf{A} \in \mathbb{R}^{n \times n}$, eigenvalues $\lambda_1, \ldots, \lambda_n$, and associated eigenvectors $\mathbf{x}_1, \ldots, \mathbf{x}_n$ the following are properties:
  3. The rank of A is equal to the number of non-zero eigenvalues of $\mathbf{A}$
  4. If $\mathbf{A}$ is non-singular then $1/\lambda_i$ is an eigenvalue of $\mathbf{A}^{-1}$ with associated vector $\mathbf{x}_i$, i.e. $\mathbf{A}^{-1}\mathbf{x}_i = (1/\lambda_i)\mathbf{x}_i$
  5. The eigenvalues of a diagonal matrix $\mathbf{D} = \text{diag}(d_1, \ldots, d_n)$ are just the diagonal entries $d_1, \ldots, d_n$

Eigenvalues and eigenvectors of symmetric matrices

- There are two important properties of eigenvalues and eigenvectors of symmetric matrices
- All eigenvalues are real
- The eigenvectors are orthonormal
- If the eigenvectors are linearly independent then we can decompose $\mathbf{A}$ as $\mathbf{U \Lambda A}^T$
- $\mathbf{U}$ is an orthogonal matrix where colums are the eigenvectors
- $\mathbf{\Lambda} = \text{diag}(\lambda_1, \ldots, \lambda_n)$