Matrix Decompositions 1 of 2

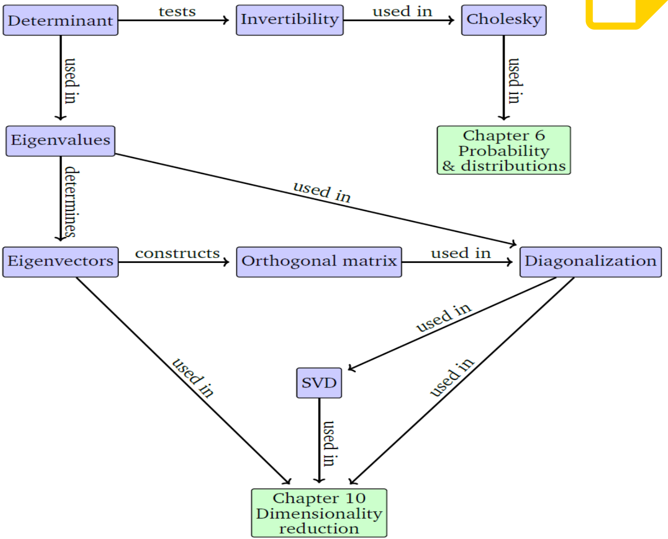
At it's core matrix decompositions is:

- The re-expression of the orginal matrix into a number of sub-components or parts.
- Typically in the form of matrix products or weighted sums of matrices.
- The re-expression typically exposed fundamental aspects of the original matrix in a more direct manner.

A simple example of this concept is in the context of natural numbers:

- For any $x \in \mathbb{N}$ we can rewrite x as a product of primes.
- For example $24 = 2 \times 2 \times 2 \times 3 = 2^3 3$

Matrix Decompositions



Source: M.P. Deisenroth *et al*, Mathematics for Machine Learning (First Edition)

Determinant

Determinants are only defined for square matrices $\mathbf{A} \in \mathbb{R}^{n \times n}$

- We denote the determinate of $\mathbf{A}$ as

$$\det(\mathbf{A}) = \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix}$$

  where

$$\det : \mathbb{R}^{n \times n} \to \mathbb{R}$$

- Before we define how to calculate the determinate the application of it is worth briefly exploring

Testing for Matrix Invertibility

Recall from Chapter 2 that in the case of a 2×2 matrix we could get an explicit formula for an inverse as follows:

- Let $\mathbf{A} \in \mathbb{R}^{2\times 2}$ then $\mathbf{A}^{-1}$ can be calculated as

$$\mathbf{A}^{-1} = \frac{1}{a_{11}a_{22} - a_{12}a_{21}} \begin{bmatrix} a_{22} & -a_{12} \\ -a_{21} & a_{11} \end{bmatrix} \tag{1}$$

- Now the inverse only exist if and only if $a_{11}a_{22} - a_{12}a_{21} \neq 0$
- This denominator is in fact $\det(\mathbf{A})$
  - So the inverse of our 2×2 matrix exist if and only if $\det(\mathbf{A}) \neq 0$

Testing for Matrix Invertibility

Theorem 4.1

For any square matrix $\mathbf{A} \in \mathbb{R}^{nxn}$ it holds that $\mathbf{A}$ is invertible if and only if $\det(\mathbf{A}) \neq 0$.

We can directly obtain the inverse for $n = 1, 2, 3$ with relative ease

- $n = 1$

$$\det(\mathbf{A}) = \det(a_{11}) = a_{11} \tag{2}$$

$ax = b$

$x = \frac{b}{a} \quad a \neq 0$

- $n = 2$

$$\det(\mathbf{A}) = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = a_{11}a_{22} - a_{12}a_{21} \tag{3}$$

- $n = 3$ Sarrus' rule

$$= a_{1} \begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix} - a_{12} \begin{vmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{vmatrix} + a_{13} \begin{vmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{vmatrix}$$

$$\det(\mathbf{A}) = \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = a_{11}a_{22}a_{33} + a_{21}a_{32}a_{13} + a_{31}a_{12}a_{23}$$

$$- a_{31}a_{22}a_{13} - a_{11}a_{32}a_{23} - a_{21}a_{12}a_{33}$$

Triangular Matrices and a Simple Determinant Formula

Triangular matrix

Let $\mathbf{T} \in \mathbb{R}^{n \times n}$ then

- $\mathbf{T}$ is upper-triangular matrix if $T_{ij}=0$ for $i > j$.
- $\mathbf{T}$ is lower-triangular matrix if $T_{ij}=0$ for $i < j$.

Upper-triangular matrix example:

$$\begin{bmatrix} 1 & 2 & 3 \\ 0 & 4 & 5 \\ 0 & 0 & 6 \end{bmatrix}$$

Lower-triangular matrix example:

$$\begin{bmatrix} 1 & 0 & 0 \\ 2 & 3 & 0 \\ 4 & 5 & 6 \end{bmatrix}$$

Triangular Matrices and a Simple Determinant Formula

If $\mathbf{T} \in \mathbb{R}^{n \times n}$ is Triangular then

$$\det(\mathbf{T}) = \prod_{i=1}^{n} T_{ii} \tag{5}$$

But what if $\mathbf{T}$ is not triangular?

- We need a generalized algorithm.
- The most common is to use the Laplace expansion which is a recursive approach.

Determinant Calculation

Laplace Expansion

Let $\mathbf{A} \in \mathbb{R}^{n \times n}$, then for all $j = 1, \ldots, n$:

1. Expansion along column j

$$\det(\mathbf{A}) = \sum_{k=1}^{n} (-1)^{k+j} a_{kj} \det(\mathbf{A}_{k,j}) \tag{6}$$

2. Expansion along row j

$$\det(\mathbf{A}) = \sum_{k=1}^{n} (-1)^{k+j} a_{jk} \det(\mathbf{A}_{j,k}) \tag{7}$$

where $\mathbf{A}_{k,j} \in \mathbb{R}^{(n-1)\times(n-1)}$ is the submatrix of $\mathbf{A}$ that we obtain when deleting row k and column j.

Determinant Calculation: Example

Consider

$$\mathbf{A} = \begin{bmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \\ 0 & 0 & 1 \end{bmatrix}$$

If we us approach 2 and select the first row we arrive at:

$$\begin{aligned} \begin{vmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \\ 0 & 0 & 1 \end{vmatrix} =& (-1)^{1+1} \cdot 1 \begin{vmatrix} 1 & 2 \\ 0 & 1 \end{vmatrix} + (-1)^{1+2} \cdot 2 \begin{vmatrix} 3 & 2 \\ 0 & 1 \end{vmatrix} + (-1)^{1+3} \cdot 3 \begin{vmatrix} 3 & 1 \\ 0 & 0 \end{vmatrix} \\ =& 1(1-0) - 2(3-0) + 3(0-0) \\ =& -5 \end{aligned}$$

Determinant Calculation: Example

Consider

$$\mathbf{A} = \begin{bmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \\ 0 & 0 & 1 \end{bmatrix}$$

We can also make our lives easy and pick the 3rd row.

$$\begin{aligned} \begin{vmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \\ 0 & 0 & 1 \end{vmatrix} =& (-1)^{3+3} 1 \cdot \begin{vmatrix} 1 & 2 \\ 3 & 1 \end{vmatrix} \\ =& 1(1-6) \\ =& -5 \end{aligned}$$

Determinant: Useful Properties and Results

Let $\mathbf{A}, \mathbf{B} \in \mathbb{R}^{n \times n}$

- $\det(\mathbf{AB}) = \det(\mathbf{A}) \det(\mathbf{B})$ *
- $\det(\mathbf{A}) = \det(\mathbf{A}^T)$
- If $\mathbf{A}$ is regular then $\det(\mathbf{A}^{-1}) = \frac{1}{\det(\mathbf{A})}$
- If $\mathbf{A}$ and $\mathbf{B}$ are similar matrices than $\det(\mathbf{A}) = \det(\mathbf{B})$
  - This means the determinant is invariant to the choice of basis of a linear mapping.

$A = S^{-1}BS$ For invert S

- Adding a multiple of a column/row to another one does not change $\det(\mathbf{A})$.

$$\det(A) = \det(S^{-1}BS) = \det(S^{-1})\det(B)\det(S) = \frac{1}{k}\det(B)\,k = \det(B)$$

- $\det(\lambda \mathbf{A}) = \lambda^n \det(\mathbf{A})$
- Swapping two rows/columns changes the sign of $\det(\mathbf{A})$.
  - Because of the last three properties, we can use Gaussian elimination to compute $\det(\mathbf{A})$ by bringing $\mathbf{A}$ into row-echelon form.

Trace

Trace

The *trace* of a square matrix $\mathbf{A} \in \mathbb{R}^{n \times n}$ is defined as

$$tr(\mathbf{A}) = \sum_{i=1}^{n} a_{ii} \tag{8}$$

The trace satisfies the following properties:

1. $tr(\mathbf{A} + \mathbf{B}) = tr(\mathbf{A}) + tr(\mathbf{B})$ for $\mathbf{A}, \mathbf{B} \in \mathbb{R}^{n \times n}$
2. $tr(\alpha \mathbf{A}) = \alpha tr(\mathbf{A})$ for $\alpha \in \mathbb{R}$, $\mathbf{A} \in \mathbb{R}^{n \times n}$
3. $tr(\mathbf{I}_n) = n$
4. $tr(\mathbf{AB}) = tr(\mathbf{BA})$ for $\mathbf{A} \in \mathbb{R}^{n \times k}$, $\mathbf{B} \in \mathbb{R}^{k \times n}$

It can be shown that only one function satisfies these four properties together– the trace (Gohberg et al., 2012).

Trace

- The fourth property actually generalizes to invariances *under cyclic permutations*, for example

$$tr(\mathbf{AKL}) = tr(\mathbf{KLA}) \tag{9}$$

for matrices $\mathbf{A} \in \mathbb{R}^{a \times k}$, $\mathbf{K} \in \mathbb{R}^{k \times l}$, $\mathbf{L} \in \mathbb{R}^{l \times a}$.

- The fourth property has a important special case namely:
  - if $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ then

$$tr(\mathbf{x}\mathbf{y}^T) = tr(\mathbf{y}^T\mathbf{x}) = \mathbf{y}^T\mathbf{x} \tag{10}$$

dot prod

$(n \times 1) \cdot (1 \times n)$

$\Rightarrow (n \times n)$

$(1 \times n) \cdot (n \times 1)$

$\Rightarrow 1$

Trace

Similar matrices and the trace

- If **A** and **B** are similiar then

$$tr(\mathbf{B}) = tr(\mathbf{S}^{-1}\mathbf{A}\mathbf{S}) = tr(\mathbf{A}\mathbf{S}\mathbf{S}^{-1}) = tr(\mathbf{A}) \tag{11}$$

  - This means that the trace does not depend on the basis used.

Characteristic Polynomial

Characteristic Polynomial

For $\lambda \in \mathbb{R}$ and a square matrix $\mathbf{A} \in \mathbb{R}^{n \times n}$

$$p_A(\lambda) := \det(\mathbf{A} - \lambda \mathbf{I}_n) \tag{12}$$

$$= c_0 + c_1\lambda + c_2\lambda^2 + \cdots + c_{n-1}\lambda^{n-1} + (-1)^n\lambda^n \tag{13}$$

where $c_0, \ldots, c_{n-1} \in \mathbb{R}$, is the the characteristic polynomial of $\mathbf{A}$. In particular,

$$c_0 = \det(\mathbf{A}) \tag{14}$$

$$c_{n-1} = (-1)^{n-1} tr(\mathbf{A}) \tag{15}$$

Eigenvalues and Eigenvectors

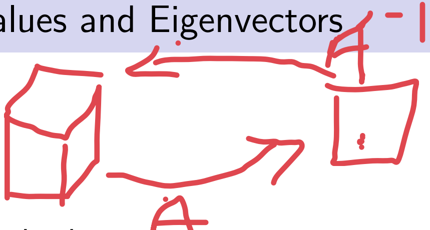
Let $\mathbf{A} \in \mathbb{R}^{n \times n}$ be a square matrix.

- Then λ is an *eigenvalue* of $\mathbf{A}$ and $\mathbf{x} \in \mathbb{R}^n \setminus \{\mathbf{0}\}$ is the corresponding *eigenvector* of $\mathbf{A}$ if

$$\mathbf{Ax} = \lambda \mathbf{x} \tag{16}$$

We call (16) the *eigenvalue equation*.

Eigenvalues and Eigenvectors



The following important statements are equivalent:

- λ is an eigenvalue of $\mathbf{A} \in \mathbb{R}^{n \times n}$
- There exists an $\mathbf{x} \in \mathbb{R}^n \setminus \{\mathbf{0}\}$ with $\mathbf{Ax} = \lambda \mathbf{x}$.
- $rk(\mathbf{A} - \lambda \mathbf{I}_n) < n$
- $\det(\mathbf{A} - \lambda \mathbf{I}_n) = 0$

$\hookrightarrow P_A(\lambda)$

$Ax = \lambda x$

$Ax = \lambda I x$

$Ax - \lambda I x = 0$

$\underline{(A - \lambda I)x} = \underline{0}$

$Tx = 0$ for $T = (A - \lambda I)$

if $\det(T) = 0$ then no unique solution (space of solns)

Collinearity and Codirection

- Two vectors that point in the same direction are called *codirected*.
- Two vectors are *collinear* if they point in the same or the opposite direction

Non-uniqueness of eigenvectors).

- $\mathbf{x}$ is an eigenvector of $\mathbf{A}$ is associated with eigenvalue λ then for any $c \in \mathbb{R} \setminus \{\mathbf{0}\}$ it holds that $c\mathbf{x}$ is an eigenvector of $\mathbf{A}$ with the same eigenvalue since

$$\mathbf{A}(c\mathbf{x}) = c\mathbf{A}\mathbf{x} = c\lambda\mathbf{x} = \lambda(c\mathbf{x}) \tag{17}$$

Thus, all vectors that are *collinear* to $\mathbf{x}$ are also eigenvectors of $\mathbf{A}$.

Eigenvalues and Eigenvectors

Theorem 4.8

$\lambda \in \mathbb{R}$ is an eigenvalue of $\mathbf{A} \in \mathbb{R}^{n \times n}$ if and only λ is a root of the characteristic polynomial $p_{\mathbf{A}}(\lambda)$ of $\mathbf{A}$

Algebraic multiplicity

Let a square matrix $\mathbf{A}$ have an eigenvalue λ_i.

- The *algebraic multiplicity* of λ_i is the number of times the root appears in the characteristic polynomial.

Eigenvalues and Eigenvectors

Eigenspace and Eigenspectrum

For $\mathbf{A} \in \mathbb{R}^{n \times n}$ the set of all eigenvectors of $\mathbf{A}$ associated with an eigenvalue λ spans a subspace ~~eigenspace~~ of $\mathbb{R}^n$,

- which is called the *eigenspace* of $\mathbf{A}$ with respect to λ and is denoted E_λ.

The set of all eigenvalues of $\mathbf{A}$ is called the *eigenspectrum*, or just *spectrum*, of $\mathbf{A}$.

$$x_{t+1} = Ax_t$$

$$E_S = \bigcup_{i=1}^{m} E_{\lambda_i}$$

Eigenvalues and Eigenvectors

For a given eigenvalue, λ, of $\mathbf{A} \in \mathbb{R}^{n \times n}$ it is worth noting that

- E_λ is the solution space to the homogeneous system of linear equations

$$(\mathbf{A} - \lambda \mathbf{I})\mathbf{x} = \mathbf{0} \tag{18}$$

Geometrically, the eigenvector corresponding to a nonzero eigenvalue, points in a direction that is **stretched** by the linear mapping.

- The eigenvalue is the factor by which it is stretched.
- If the eigenvalue is negative, the direction of the stretching is flipped.

The Case of the Identity Matrix

A simple but interesting case worth considering is that of $\mathbf{I} \in \mathbb{R}^{n \times n}$:

- Note that

$$p_{\mathbf{I}}(\lambda) = \det(\mathbf{I} - \lambda\mathbf{I}) \tag{19}$$
$$(1 - \lambda)^n \tag{20}$$

  In order to find the eigenvalues we set $p_{\mathbf{I}}(\lambda) = 0$ as such
  - The only solution is $\lambda = 1$ which will have multiplicity of n
- Moreover, $\mathbf{I}\mathbf{x} = \lambda\mathbf{x} = 1\mathbf{x}$ holds for all $\mathbf{x} \in \mathbb{R}^n \setminus \{\mathbf{0}\}$
  - This means that E_1 of the identity matrix spans n dimensions,
  - and all n standard basis vectors of $\mathbb{R}^n$ are eigenvectors of $\mathbf{I}$.

Eigenvalues and Eigenvectors: Useful Properties

Commonly used properties of eigenvalues and eigenvector

- A matrix $\mathbf{A}$ and its transpose $\mathbf{A}^T$
  - but **not necessarily** the same eigenvectors.
- The eigenspace E_λ is the null space of $\mathbf{A} - \lambda \mathbf{I}$ since

$$\begin{aligned} \mathbf{A}\mathbf{x} = \lambda \mathbf{x} &\iff \mathbf{A}\mathbf{x} - \lambda \mathbf{x} = 0 &(21)\\ &\iff (\mathbf{A} - \lambda \mathbf{I})\mathbf{x} = 0 &(22)\\ &\iff \mathbf{x} \in ker(\mathbf{A} - \lambda \mathbf{I}) &(23) \end{aligned}$$

Eigenvalues and Eigenvectors: Useful Properties

$A = S^{-1}BS$

Commonly used properties of eigenvalues and eigenvector

- Similar matrices possess the same eigenvalues.
  - Therefore, a linear mapping Φ has eigenvalues that are independent of the choice of basis of its transformation matrix.
  - So we know have three key characteristic parameters of a matrix that are independent of the choice of basis:
    - eigenvalues
    - determinant
    - trace
- Symmetric, positive definite matrices always have positive, real eigenvalues.