Linear Algebra

Prof. Ioannis Pitas
Aristotle University of Thessaloniki
pitas@csd.auth.gr
www.aiia.csd.auth.gr
version 3.0

Artificial Intelligence &
Information Analysis Lab

Linear Algebra

- **Vectors, matrices**
- System of linear equations
- Eigenanalysis
- Singular value Decomposition
- Other matrix decompositions
- Tensors Fundamentals
- BLAS.

Vectors

A ***vector*** of dimension n is an 1D array of numbers:

$$\mathbf{x} = [x_1, \dots, x_n]^T \in \mathbb{R}^n.$$

- $x_1, \dots, x_n$: n vector coordinates.

- ***Vector inner product***:

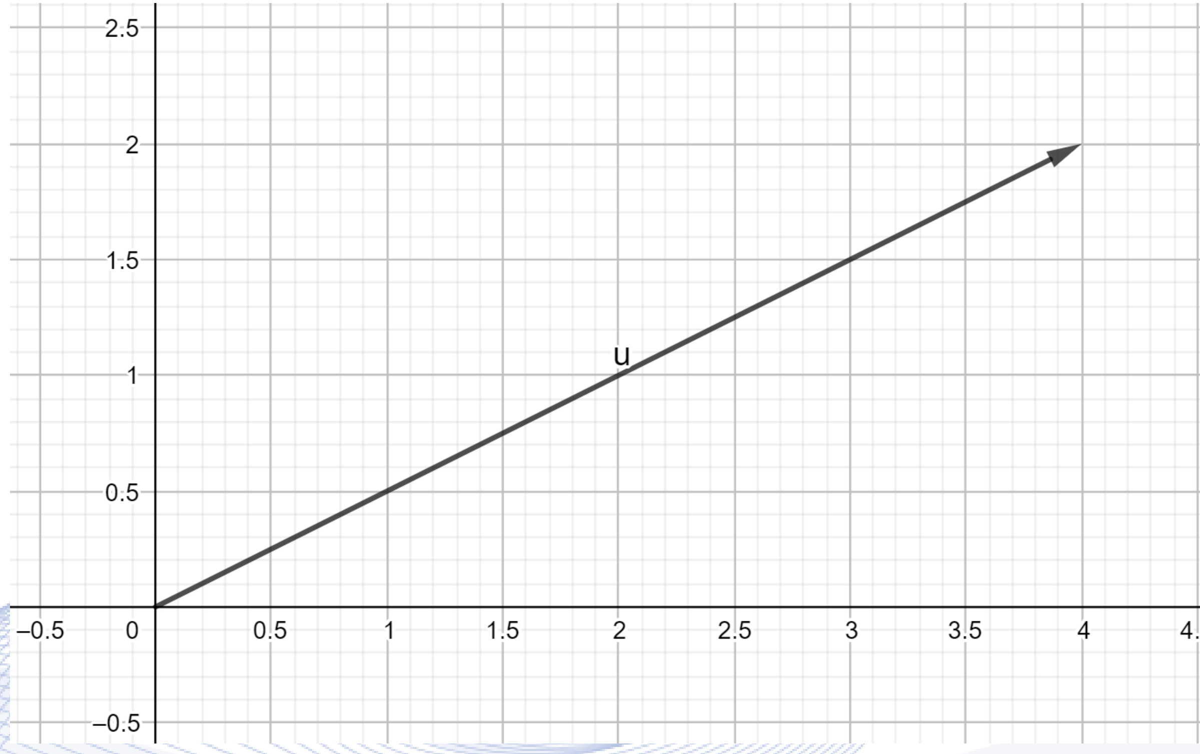
$$\mathbf{x}^T\mathbf{y} = \sum_{k=1}^{n} x_k y_k.$$

Artificial Intelligence & Information Analysis Lab

Vectors

Geometrical vector interpretation:

- a point in a Euclidean space $\mathbb{R}^n$.
- ***Polar, spherical vector representation*** in $\mathbb{R}^2$, $\mathbb{R}^3$: magnitude $|\mathbf{x}|$ and direction angle(s).



Cartecian vector $\boldsymbol{x} = [4,2]^T$ representation.

Matrices and tensors

A **matrix** $\mathbf{A}$ is a $n \times m$ table (2D array) of numbers:

$$\mathbf{A} = \left[a_{ij}\right] = [\mathbf{a}_1 \ldots \mathbf{a}_m] = \begin{bmatrix} \alpha_{11} & \alpha_{12} & \cdots & \alpha_{1m} \\ \alpha_{21} & \alpha_{22} & \cdots & \alpha_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ \alpha_{n1} & \alpha_{n2} & \cdots & \alpha_{nm} \end{bmatrix}.$$

- $\mathbf{a}_j, j = 1, \ldots, m$: matrix columns.
- k–th order ***tensor***: A k-D array of numbers $\mathbf{A} \in \mathbb{R}^{n_1 \times \cdots \times n_k}$.

Matrix Properties

- The multiplication of ***inverse matrix*** $\mathbf{A}^{-1}$ of a square matrix $\mathbf{A} \in \mathbb{R}^{n \times n}$ with matrix $\mathbf{A}$ is the identity $n \times n$ matrix $\mathbf{I}$:

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

- It exists, if matrix determinant $\det(\mathbf{A}) \neq 0$.
- Inverse matrix properties:

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

- A square matrix that is not invertible is called ***singular***.
- $\mathbf{A}$ is ***singular***, if its rank is less than n.

Linear Algebra

- Vectors, matrices
- **System of linear equations**
- Eigenanalysis
- Singular value Decomposition
- Other matrix decompositions
- Tensors Fundamentals
- BLAS.

System of linear equations

System of linear equations

$$\mathbf{Ax} = \mathbf{y}.$$

- Unknown vector: $\mathbf{x} \in \mathbb{R}^m$.
- m equations:

$$\begin{bmatrix} y_1 \\ . \\ . \\ . \\ y_n \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1m} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nm} \end{bmatrix} \begin{bmatrix} x_1 \\ . \\ . \\ . \\ x_m \end{bmatrix},$$

Artificial Intelligence &
Information Analysis Lab

System of linear equations

System of linear equations

$$\mathbf{Ax} = \mathbf{y}.$$

- If $m = n$ and $\mathbf{A}^{-1}$ exists, the solution is:

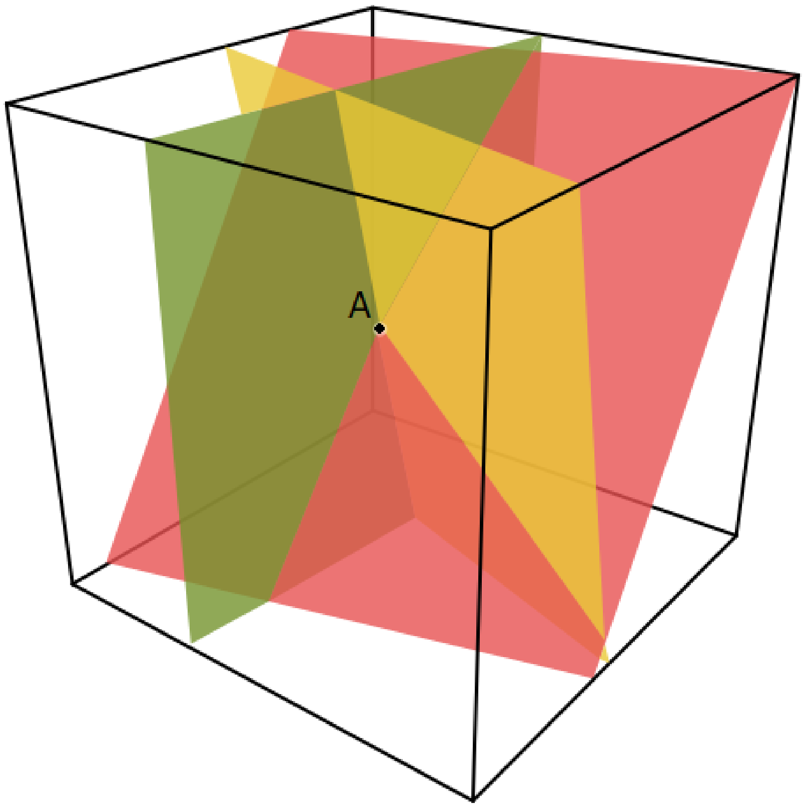
$$\mathbf{x} = \mathbf{A}^{-1}\mathbf{y}.$$

- If $m > n$, the system is ***over-determined***.
  - Possibly no solution.
- If $m < n$, the system is ***under-determined***.
  - Possibly multiple solutions.

System of linear equations

Linear system of three variables:

- Each equation determines a plane in space $\mathbb{R}^3$.
- Their intersection point is the solution of the system.



Linear Algebra

- Vectors, matrices
- System of linear equations
- **Eigenanalysis**
- Singular value Decomposition
- Other matrix decompositions
- Tensors Fundamentals
- BLAS.

Eigenanalysis

- Eigenvectors and eigenvalues of a matrix are important as they provide fundamental information about a matrix.
- They allow easy determination as to whether a matrix is positive definite or not.
- Also allow determination as to whether a matrix is invertible and how sensitive to numerical errors the inverse will be.

Eigenanalysis

If $\mathbf{A}$ is an $n \times n$ matrix, its eigenvalue λ and eigenvector $\mathbf{v}$ satisfy:

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

Equivalently, they form a solution of the homogeneous linear equation system:

$$(\mathbf{A} - \lambda\mathbf{I}_n)\mathbf{v} = 0,$$

where $\mathbf{I}_n$ is a unitary $n \times n$ matrix.

Spectral Theorem

Matrix $\mathbf{A}$ is ***guaranteed*** to be invertible, if $\lambda_i \neq 0$ for all $i = 1, \dots, n$.

This is equivalent to matrix determinant being not equal to 0:

$$\det(\mathbf{A}) = \Pi_{i=1}^{n} \lambda_i \neq 0.$$

- In practical Machine Learning and Computer Vision applications, matrices are estimated from sample data.
- Therefore may be ***ill-conditioned***, if one or more of the eigenvalues are close to zero.

Positive definite matrices

Positive definite matrix definition:

$$\mathbf{x}^T\mathbf{C}\mathbf{x} > 0, \qquad \forall \mathbf{x} \in \mathbb{R}^n.$$

for symmetric matrices $\mathbf{C}$.

- Eigenvalues of a positive definite matrix are positive:

$$\lambda_i > 0, i = 1, \dots, n.$$

- Determinant of a positive definite matrix:

$$\det(\mathbf{C}) = \Pi_{i=1}^{n} \lambda_i > 0.$$

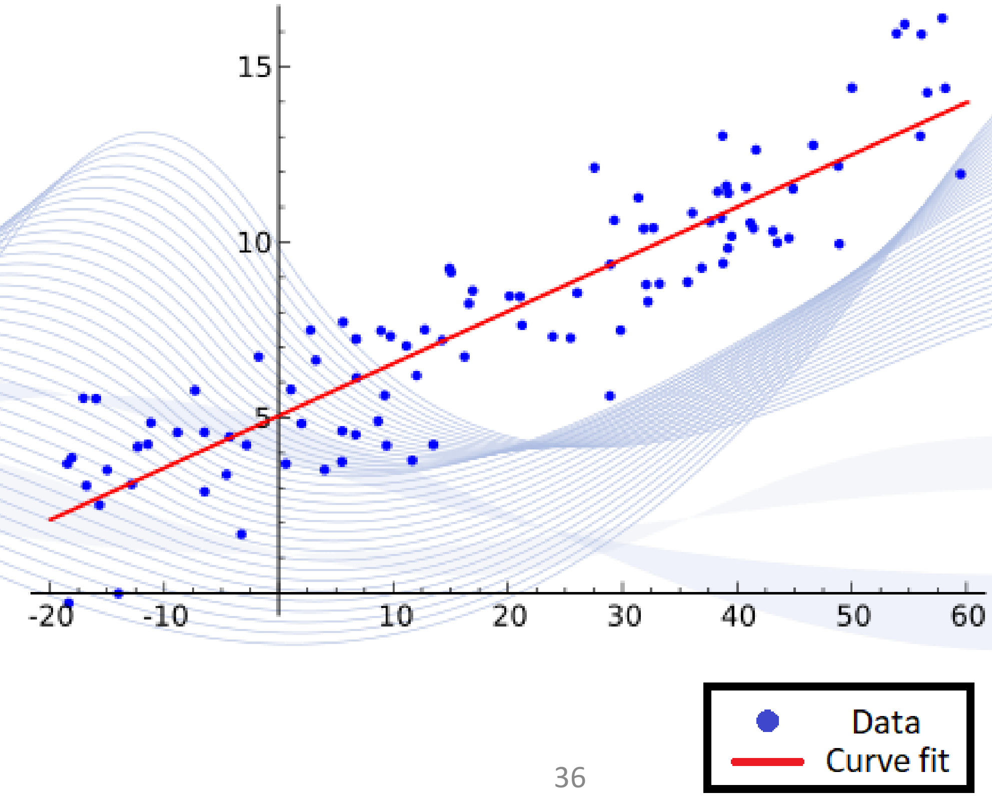
Linear Algebra

- Vectors, matrices
- System of linear equations
- Eigenanalysis
- **Singular value Decomposition**
- Other matrix decompositions
- Tensors Fundamentals
- BLAS

Singular Value Decomposition

Singular Value Decomposition (***SVD***) deals with:

- Systems of equations whose matrices are singular or numerically very close to singular.
- Solving most ***Linear Least-Squares*** (***LLS***) problems.
- Providing low rank matrix approximations.



Artificial Intelligence & Information Analysis Lab

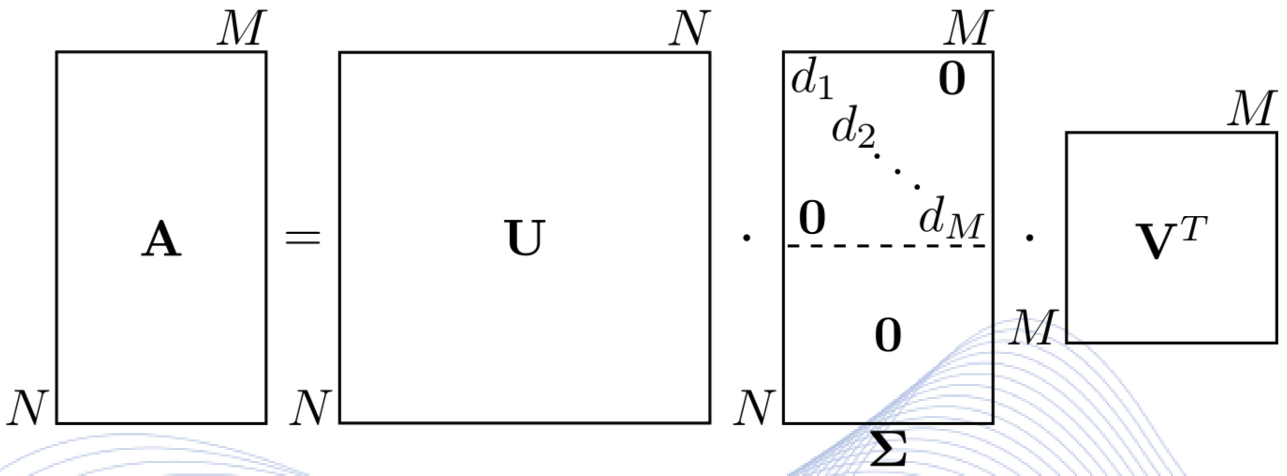
Singular Value Decomposition

Any $n \times m$ matrix $\mathbf{A}$ can be decomposed into:

$$\mathbf{A} = \mathbf{U}\boldsymbol{\Sigma}\mathbf{V}^T,$$

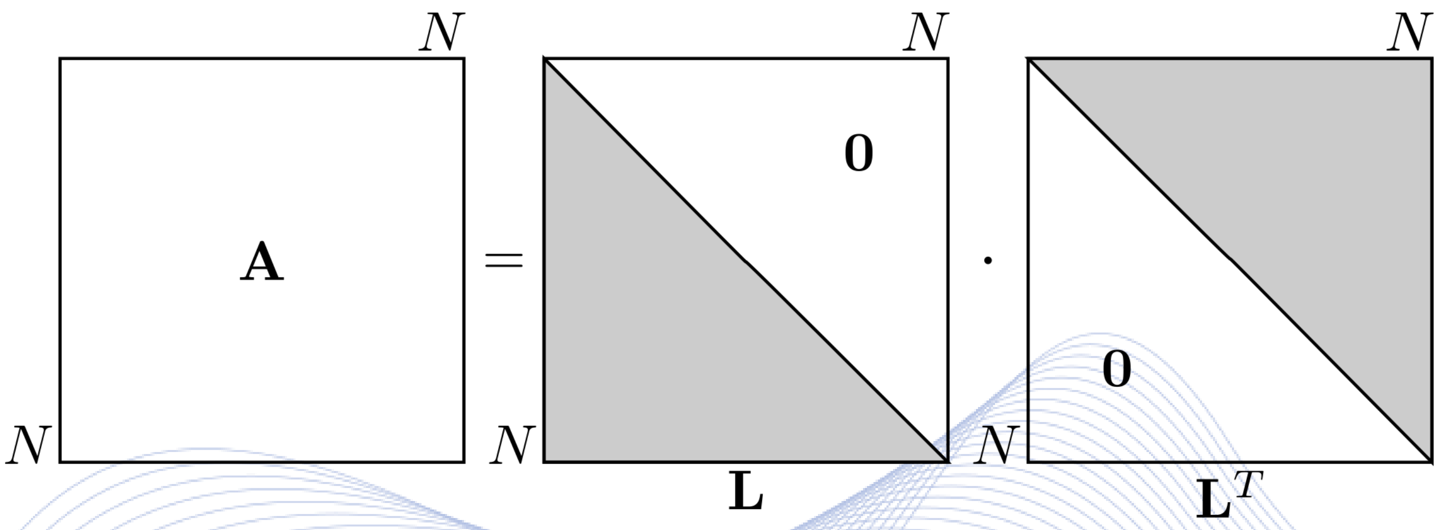
- $\mathbf{U}$ ($n \times n$ orthogonal) unitary matrix,
- $\boldsymbol{\Sigma}$ ($n \times m$ diagonal) matrix and
- $\mathbf{V}^T$ ($m \times m$ orthogonal) unitary matrix.
- ***Singular values*** of $\mathbf{A}$ are the $r = \min(n, m)$ $\sigma_1, \sigma_2, \dots, \sigma_r$ diagonal elements of $\boldsymbol{\Sigma}$.

Singular Value Decomposition



Singular Value Matrix Decomposition.

Cholesky decomposition



Cholesky matrix decomposition.

Cholesky decomposition

Cholesky matrix decomposition is mainly used for the numerical solution of linear equations $\mathbf{Ax} = \mathbf{b}$, when $\mathbf{A}$ is ***Symmetric Positive Definite*** (***SPD***) matrix.

- We first compute the matrix $\mathbf{L}$ as described above. Then, we solve the equation:

$$\mathbf{Ly} = \mathbf{b},$$

where $\mathbf{y} = \mathbf{L}^T\mathbf{x}$, using forward substitution and, finally, we solve:

$$\mathbf{L}^T\mathbf{x} = \mathbf{b},$$

using back substitution.

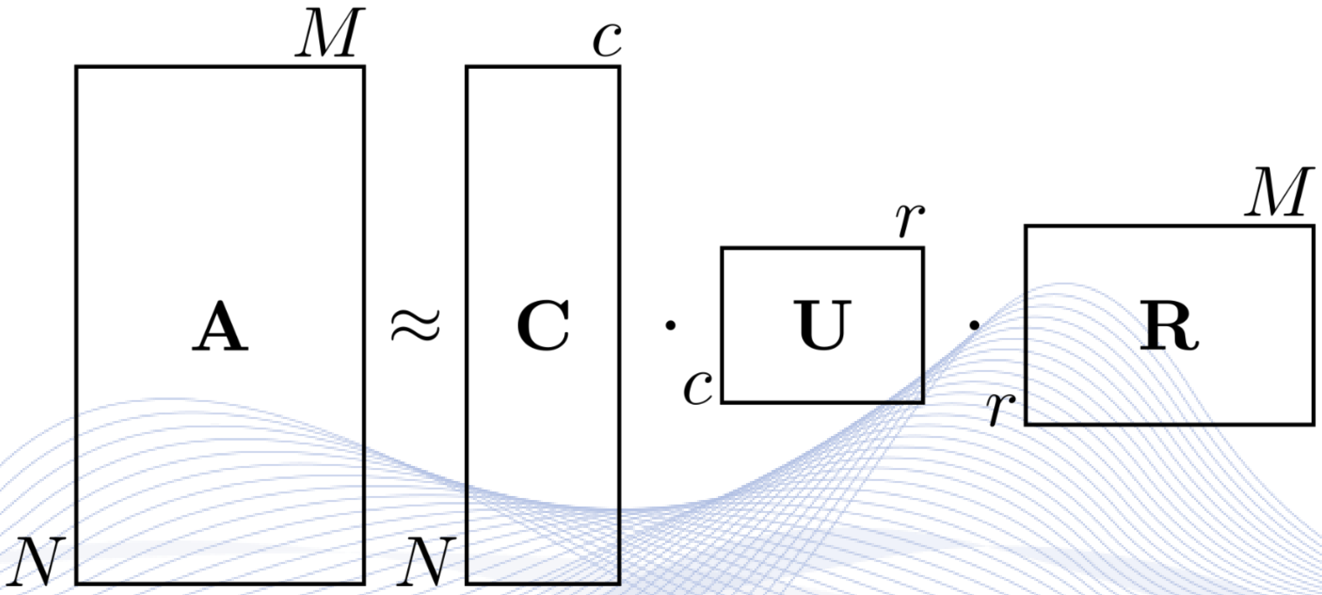
Artificial Intelligence & Information Analysis Lab

CUR Matrix Decomposition

- In ***CUR matrix approximation***, the multiplication of three matrices $\mathbf{C} \in \mathbb{R}^{m \times c}, \mathbf{U} \in \mathbb{R}^{c \times r}, \mathbf{R} \in \mathbb{R}^{r \times n}$ closely approximate a given matrix $\mathbf{A}$, by minimizing the approximation error $||\mathbf{A} - \mathbf{CUR}||_F$.

- A CUR approximation can be used as low-rank approximation.

CUR Matrix Decomposition



CUR matrix approximation.

Non-negative matrix factorization

- Data matrix $\mathbf{X}$ is an $n \times N$ matrix containing N data vectors $\mathbf{X} = [\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_N]$.
- It can be decomposed in a product of $n \times p$ and $p \times N$ matrices $\mathbf{F}$ and $\mathbf{H}$, respectively:

$$\mathbf{X} = \mathbf{F}\mathbf{H}.$$

- p is smaller than N and n.
- All elements of matrices $\mathbf{F}, \mathbf{H}$ should be positive: $f_{ij} \geq 0$, $h_{kl} \geq 0$.

Non-negative matrix factorization

$\mathbf{x}_i =$  $\approx h_{1i}$  $+h_{2i}$  $+ \cdots h_{li} +$  $+h_{pi}$ 

NMF image decomposition.

Other matrix decompositions

Some other matrix decompositions are:

- ***Polar decomposition***, applicable to a square, complex matrix $\mathbf{A}$: $\mathbf{A} = \mathbf{UP}$ or $\mathbf{A} = \mathbf{P}'\mathbf{U}$.
- ***Mostow decomposition***, applicable to a square, complex, non-singular matrix $\mathbf{A}$: $\mathbf{A} = \mathbf{U}e^{i\mathbf{M}}e^{\mathbf{S}}$.
- ***Sinkhorn normal form***, applicable to a square, real matrix $\mathbf{A}$ with strictly positive elements $\mathbf{A} = \mathbf{D}_1\mathbf{S}\mathbf{D}_2$.

There are many more matrix decompositions.

Linear Algebra

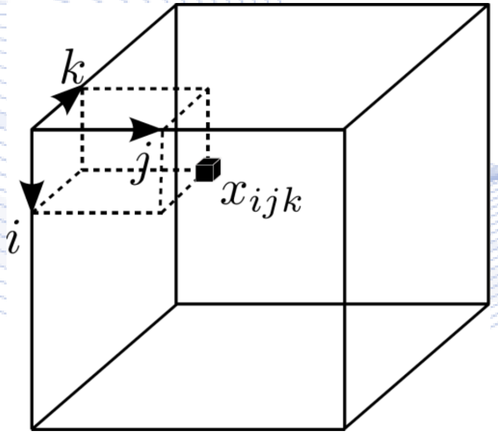
- Vectors, matrices
- System of linear equations
- Eigenanalysis
- Singular value Decomposition
- Other matrix decompositions
- **Tensor Fundamentals**
- BLAS

Tensor Fundamentals

- A ***tensor*** is a multidimensional array of numerical values, whose elements are identified using multiple indices.
- The order (degree) of a tensor is equal to the dimensionality of its array.
- Tensors can be considered to be a generalization of scalars, vectors and matrices.
  - A scalar $x \in \mathbb{R}$is a 0^{th} order tensor, a vector $\mathbf{x} \in \mathbb{R}^{n_1}$ is a 1^{st} order tensor and a matrix $\mathbf{A} \in \mathbb{R}^{n_1 \times n_2}$ is a 2^{nd} order tensor.

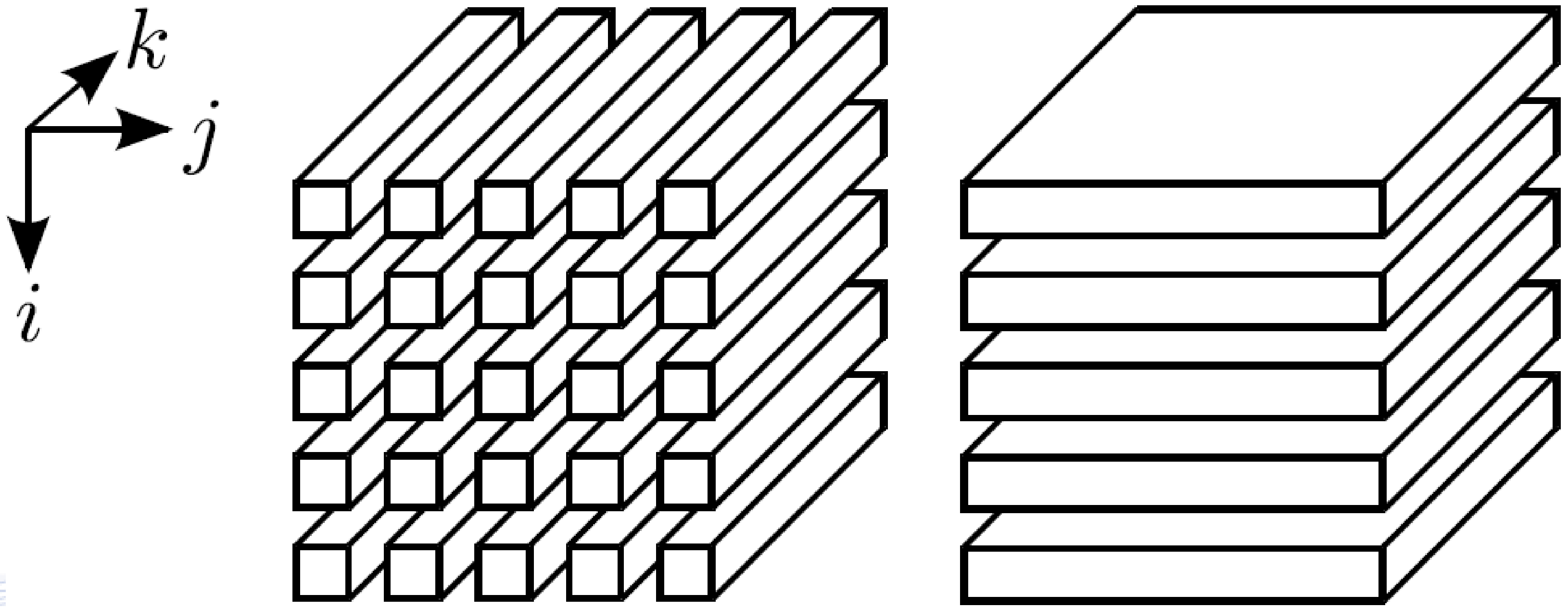
Tensor Fundamentals

- Color images and grayscale videos can be represented as 3D matrices (3^{rd} order tensors), while color videos can be represented as 4D matrices (4^{th} order tensors).
- In social media, tensors can be used to represent hypergraphs and multigraphs.



3^{rd} order tensor.

Tensor Fundamentals



Tube fibers $X_{(:jk)}$ of 3rd order tensor.

Horizontal slices $X_{(i::)}$ of a 3rd order tensor.

Artificial Intelligence & Information Analysis Lab

Tensor Fundamentals

- The Frobenius norm of a k-th order $n_1 \times n_2 \times \cdots \times n_k$ tensor $\mathbf{X} \in \mathbb{R}^{n_1 \times n_2 \times \cdots \times n_k}$ is defined as:

$$\|\mathbf{X}\|_F \triangleq \sqrt{\sum_{i_1=1}^{n_1} \sum_{i_2=1}^{n_2} \dots \sum_{i_k=1}^{n_k} X_{i_1 i_2 \dots i_k}^2}.$$

It can be used to measure the distance between tensors $\mathbf{X}$ and $\mathbf{Y}$:

$$d(X - Y) = \|\mathbf{X} - \mathbf{Y}\|_F = \sqrt{\sum_{i_1=1}^{n_1} \sum_{i_2=1}^{n_2} \dots \sum_{i_k=1}^{n_k} (X_{i_1 i_2 \dots i_k} - Y_{i_1 i_2 \dots i_k})^2}.$$

Tensor Fundamentals

- The ***inner product*** $z \triangleq \langle \mathbf{X}, \mathbf{Y} \rangle$ of the k –th order $n_1 \times n_2 \times \cdots \times n_k$ tensors $\mathbf{X}, \mathbf{Y}$ is a scalar value define as:

$$z \triangleq \sum_{i_1=1}^{n_1} \sum_{i_2=1}^{n_2} \dots \sum_{i_k=1}^{n_k} X_{i_1 i_2 \dots i_k} Y_{i_1 i_2 \dots i_k}.$$

- The inner product $\mathbf{Z} \triangleq \langle \mathbf{X}, \mathbf{Y} \rangle_{p,q}$ of a k-th order tensor $\mathbf{X}$ and a l-th order tensor $\mathbf{Y}$ ($n_p = m_q = U$) is a $(k + l - 1)$-th order $n_1 \times n_2 \times \cdots \times n_{p-1} \times n_{p+1} \times \cdots \times n_K \times m_1 \times m_2 \times \cdots \times m_{q-1} \times m_{q+1} \times \cdots \times m_l$ tensor:

$$z_{i_1 i_2 \dots i_{p-1} i_{p+1} \dots i_k j_1 j_2 \dots j_{q-1} j_{q+1} \dots j_l} \triangleq \sum_{u=1}^{U} x_{i_1 i_2 \dots i_{p-1} u i_{p+1} \dots i_k} y_{i_1 i_2 \dots j_{q-1} \dots u j_{q+1} \dots i_k}.$$

Tensor Fundamentals

- The m –mode $\mathbf{Z} \triangleq \mathbf{X} \times_m \mathbf{Y}$ product of k-th order tensor $\mathbf{X}$ and a $m \times n_m$ matrix $\mathbf{Y}$ is a k-th order $n_1 \times n_2 \times \cdots \times n_{M-1} \times m \times n_{m+1} \times \cdots \times n_k$ tensor defined as:

$$\mathbf{Z} = \langle \boldsymbol{X}, \mathbf{Y} \rangle_{m,2}.$$

- The outer product $\mathbf{Z} \triangleq \mathbf{x} \otimes \mathbf{y}$ of a m – dimensional vector $\mathbf{x}$ with a n-dimensional vector $\mathbf{y}$ is a $m \times n$ matrix, given by:

$$z_{ij} = x_i y_j.$$

Artificial Intelligence & Information Analysis Lab

Tensor Fundamentals

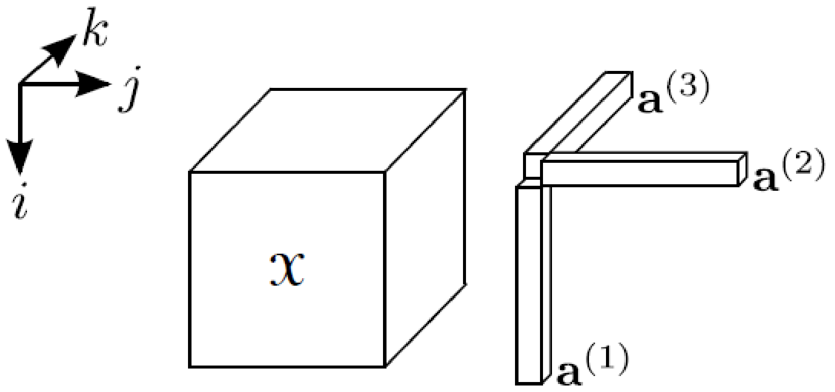
- The ***outer product*** $\mathbf{Z} \triangleq \mathbf{x} \otimes \mathbf{Y}$ of a m-dimensional vector $\mathbf{x}$ and a $k-$th order tensor $\mathbf{Y}$, is a $(k+1)-$th order tensor, defined by multiplying all the elements of $\boldsymbol{Y}$ with each element of $\mathbf{x}$:

$$z_{i j_i j_2 \ldots j_k} = x_i y_{j_1 j_2 \ldots j_k}.$$

- The outer product $\mathbf{Z} \triangleq \mathbf{X} \otimes \mathbf{Y}$ of a $k-$th order tensor $\boldsymbol{X}$ and a $l-$th order tensor $\boldsymbol{Y}$ is a $(k+l)-$th order $n_1 \times n_2 \times \cdots \times n_k \times m_1 \times m_2 \times \cdots \times m_l$ tensor:

$$z_{i_i i_2 \ldots i_k j_i j_2 \ldots j_l} = x_{i_1 i_2 \ldots i_k} y_{j_i j_2 \ldots j_l}.$$

Tensor Fundamentals



Rank 1 tensor.

Tensor Fundamentals

- The rank of a tensor $r(\mathbf{X})$ is the smallest integer r, which indicates the number of the rank-one tensors, whose sum generates $\boldsymbol{X}$:

$$\mathbf{X} = \sum_{i=1}^{r} \mathbf{B}_i \,,$$

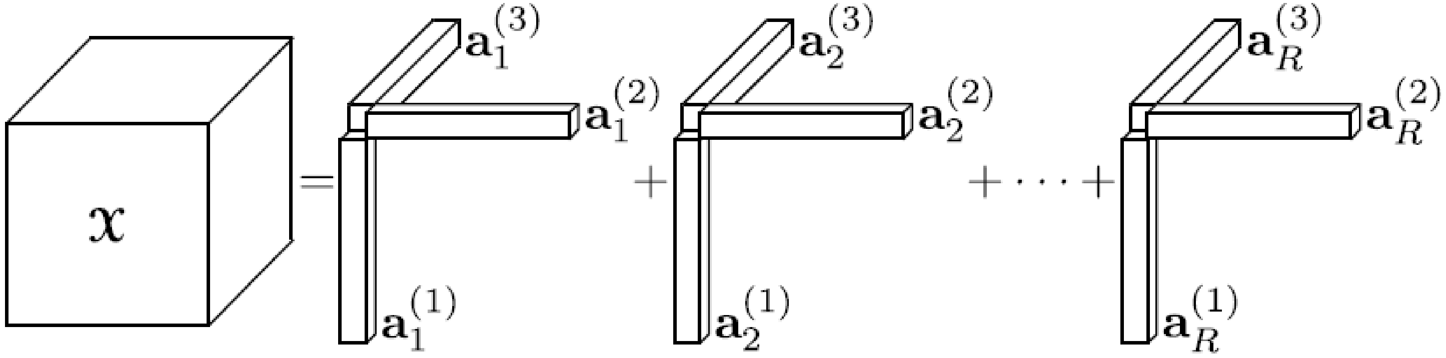
- $\mathbf{B}_i, \mathrm{i} = 1,2, \dots, r$ are the rank-one tensors.
- This is a rank r decomposition of tensor $\mathbf{X}$.
- The two main types of tensor decomposition are PARAFAC and Tucker decomposition.

Tensor Fundamentals

Parallel Factor Analysis (***PARAFAC***) decomposes a $k-$th order tensor into a sum of R rank-one tensors each composed of k linear components:

$$\mathbf{X} = \sum_{r=1}^{R} \lambda_r \mathbf{a}_r^{(1)} \times \mathbf{a}_r^{(2)} \times \cdots \times \mathbf{a}_r^{(k)}, \qquad \mathbf{a}_r^{(1)} \in \mathbb{R}^{n_1}, \mathbf{a}_r^{(2)} \in \mathbb{R}^{n_2}, \ldots, \mathbf{a}_r^{(k)} \in \mathbb{R}^{n_k}.$$

- λ_r is a factor scaling the contribution of the $r-$th rank-one tensor.

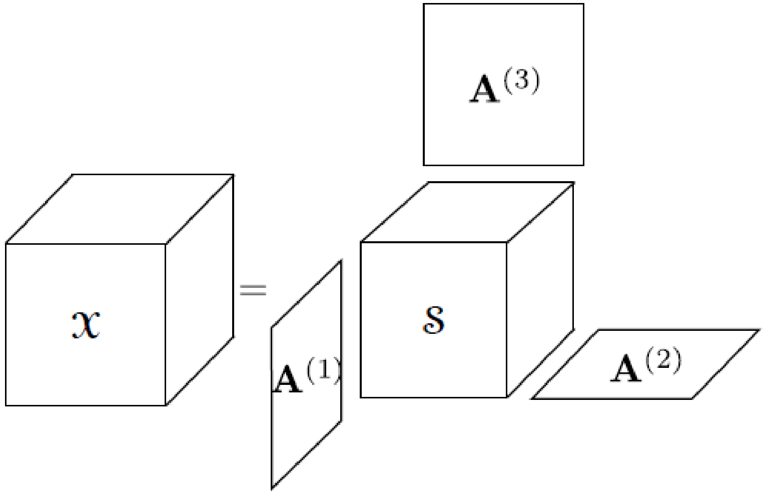


PARAFAC tensor decomposition.

Tensor Fundamentals

Tucker decomposition decomposes a $k-$th order tensor into mode products of a core tensor $\mathbf{S} \in \mathbb{R}^{m_1 \times m_2 \times \cdots m_k}$ and k matrices $\mathbf{A}^{(i)} \in \mathbb{R}^{n_i \times m_i}, i = 1, \dots, k$, each of them corresponding to a mode of $\mathbf{X}$:

$$\mathbf{X} = \mathbf{S} \times_1 \mathbf{A}^{(1)} \times_2 \mathbf{A}^{(2)} \dots \times_k \mathbf{A}^{(k)}.$$



Tucker tensor decomposition.

Linear Algebra

- Vectors, matrices
- System of linear equations
- Eigenanalysis
- Singular value Decomposition
- Other matrix decompositions
- Tensor Fundamentals
- **BLAS**

Basic Linear Algebra Subprograms (BLAS) Library

Basic Linear Algebra Subprograms (BLAS) is a software library of high-performance Linear Algebra routines.

- BLAS has three routine sets (“levels“).
- They correspond to both the chronological order of definition and publication, as well as the degree of algorithm complexity.

Q & A

Thank you very much for your attention!

More material in
http://icarus.csd.auth.gr/cvml-web-lecture-series/

Contact: Prof. I. Pitas
pitas@csd.auth.gr