From Statistics to Data Mining

Master 1
COlour in Science and Industry (COSI)
Cyber-Physical Social System (CPS2)
Saint-Étienne, France

Fabrice MUHLENBACH
https://perso.univ-st-etienne.fr/muhlfabr/
e-mail: fabrice.muhlenbach@univ-st-etienne.fr

- Introduction

➢ **Linear algebra →**
branch of mathematics concerning vector spaces, linear transformations, systems of linear equations...

➢ **Interests**

- solving so-called “linear” equations
- linear maps in vector space
→ representation of shifting in elementary geometric spaces such as a straight line, plane or physical space
→ generalization of the notion of space to any dimensions

Linear Algebra

- Introduction

➢ **History**

- Al-Khawarizmi (780-850) → « algebra »
  → translation of Indian mathematics texts and reinterpretation of Greek school work
- René Descartes (1596-1650) → association between geometry and algebra thanks to the notion of coordinates ("Cartesian" coordinate system)
- Carl Friedrich Gauss (1777-1855)
  → generic method for solving systems of linear equations ("Gaussian elimination" = row reduction)
  → linear algebra becomes a branch of mathematics in its own right

Linear Algebra

- Introduction

➢ Basic notions

- an **eigenvector** is a linear map of a space in itself
- an eigenvector corresponds to the study of the privileged axes, according to which the mapping behaves like a **dilation**, multiplying the vectors by the same constant
- this expansion ratio is called **eigenvalue**, the vectors to which it applies are called **eigenvectors**, united in an **eigen space**
- knowing the eigenvectors and eigenvalues provides key information on the linear map considered

Linear Algebra

- Introduction

➢ **Some examples of linear algebra applications:**

- geometry → study of quadratic shapes
- functional analysis
- classical mechanics → various problems (e.g. study of the movements of a vibrating string)
- quantum mechanics → study of the Schrödinger equation
- general theory of relativity → determining the space-time structure
- convex optimization → search for minima
- *Google* → web page ranking algorithm
- linear regression
- principal component analysis
- k-means classification: intra / inter-class variance optimizat°

Linear Algebra

- Introductory example: linear equation system

- solving the following system of linear equations:

$$(S) = \begin{cases} 3x_1 & +2x_2 & = 7 \\ x_1 & -3x_2 & = -5 \end{cases}$$

- this system of linear equations S can be solved in different ways (e.g., by Gaussian elimination, by the Gauss-Jordan method, etc.)
- possible use of matrix notation as an approach
- for that, we rewrite S as follows:

$$(S): AX = b$$

- where $A = \begin{pmatrix} 3 & 2 \\ 1 & -3 \end{pmatrix}$, $X = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$ and $b = \begin{pmatrix} 7 \\ -5 \end{pmatrix}$
- **theorem:** if A is a non singular matrix, then the equation system $AX = b$ has the solution $X = A^{-1}b$

Linear Algebra

- Introductory example: linear equation system

○ linear equation system:

$$(S) = \begin{cases} 3x_1 & +2x_2 & = 7 \\ x_1 & -3x_2 & = -5 \end{cases}$$

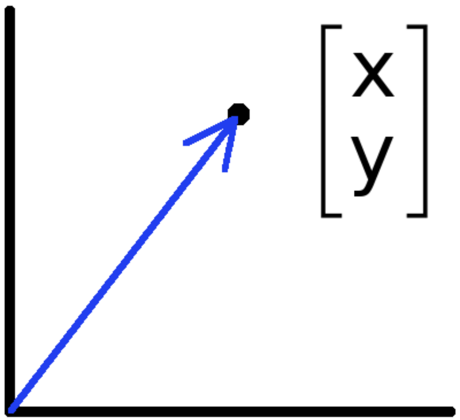
○ rewriting of S: $(S): AX = b$
where $A = \begin{pmatrix} 3 & 2 \\ 1 & -3 \end{pmatrix}$, $X = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$ and $b = \begin{pmatrix} 7 \\ -5 \end{pmatrix}$

○ solution $X = A^{-1}b$ (need to compute the inverse matrix of A)

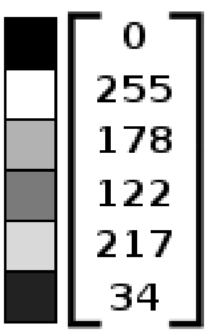
○ $X = A^{-1}b = \begin{pmatrix} \frac{3}{11} & \frac{2}{11} \\ \frac{1}{11} & \frac{-3}{11} \end{pmatrix} \begin{pmatrix} 7 \\ -5 \end{pmatrix} = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$

Linear Algebra

- Introduction: linear algebra and image processing



- vectors can represent an offset in 2D or 3D space
- points are just vectors from the origin



- data (pixels, gradients at an image keypoint, etc.) can also be treated as a vector
- such vectors do not have a geometric interpretation, but calculations like “distance” can still have value

Linear Algebra

- Introduction: linear algebra and image processing

- a matrix $A \in \mathbb{R}^{m \times n}$ is an array on numbers with size m by n, i.e., m rows and n columns

$$A = \begin{bmatrix} a_{1;1} & \cdots & a_{1;n} \\ \vdots & \ddots & \vdots \\ a_{m;1} & \cdots & a_{m;n} \end{bmatrix}$$

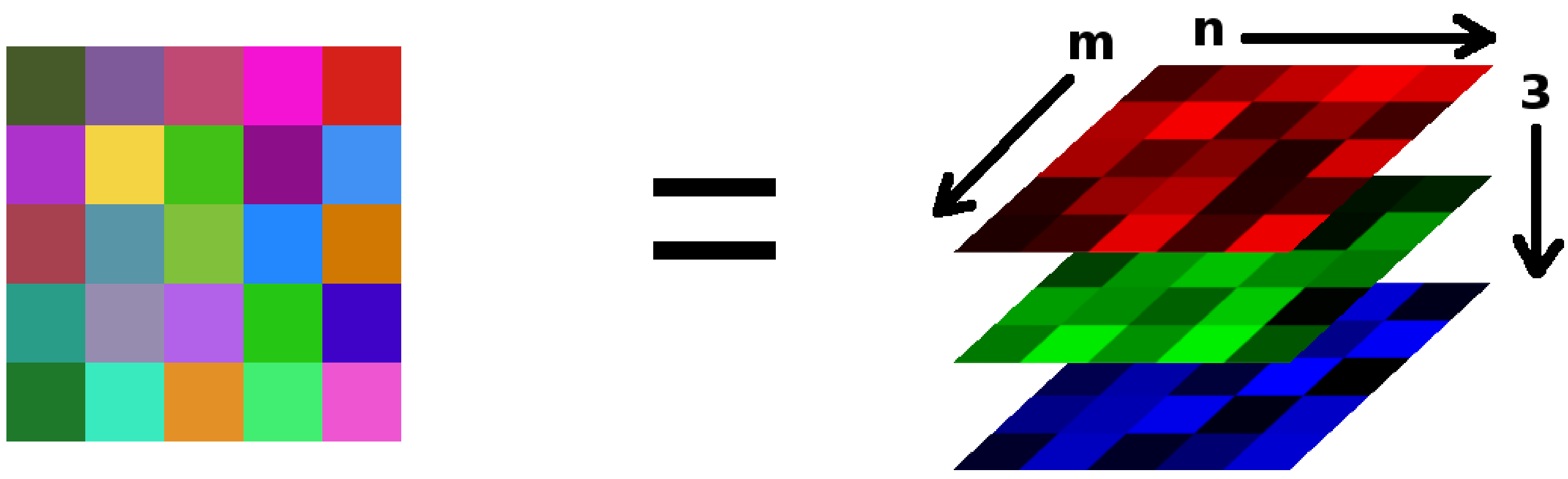
$$= \begin{bmatrix} 193 & 180 & 210 & 112 & 125 \\ 189 & 8 & 177 & 97 & 114 \\ 100 & 71 & 81 & 195 & 165 \\ 167 & 12 & 242 & 203 & 181 \\ 44 & 25 & 9 & 48 & 192 \end{bmatrix}$$

- a digital image is represented as a matrix of pixel brightness
- note that, in many computer languages, the upper left corner is [y,x] = (0,0)

Linear Algebra

- Introduction: linear algebra and image processing

  - grayscale images have one number per pixel, and are stored as an $m \times n$ matrix
  - color images have 3 numbers per pixel –red, green, and blue brightnesses (RGB)
  - color matrices are stored as $m \times n \times 3$ matrices



Linear Algebra

- Matrix Multiplication

➢ Definition

- matrix multiplication
  = binary operation that produces a matrix from two matrices
- a matrix A is characterized by its dimensions: m rows et n columns ($m \times n$ matrix) and can be multiplied with a matrix B with dimension ($m' \times n'$ matrix) iff $n = m'$
  → for matrix multiplication, the number of columns in the first matrix (n) must be equal to the number of rows in the second matrix (m')
- resulting matrix = **matrix product**, has the number of rows of the first and the number of columns of the second matrix
  → the matrix product AB has dimensions ($m \times n'$)

- Matrix Multiplication

➢ **Definition**

- if the (i, k)-entry of a $m \times n$ matrix A is indicated by a_{ik}, and the (k, j)-entry of a $n \times p$ matrix B is indicated by b_{kj}, then the (i, j)-entry of the $m \times p$ product matrix, denoted AB, is indicated by c_{ij}, is the sum of the products of corresponding entries from row i of A and column j of B:

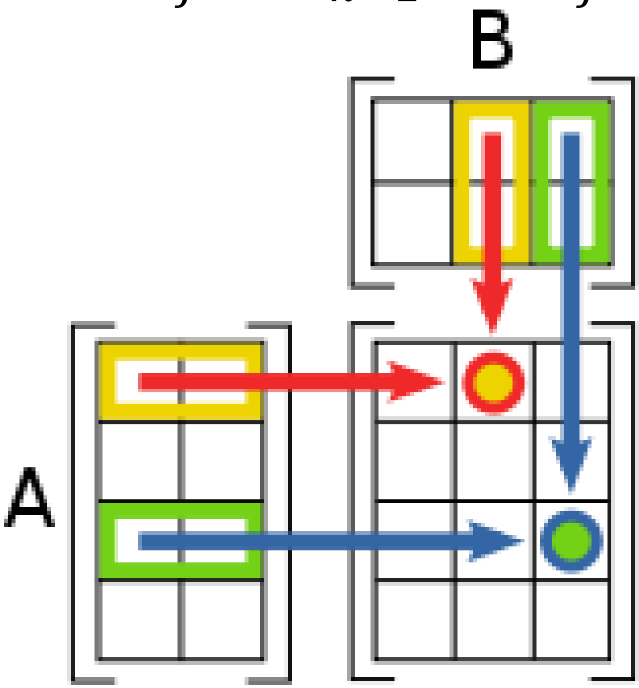
$$\forall i, j\colon\ c_{ij} = \sum_{k=1}^{n} a_{ik} b_{kj} = a_{i1} b_{1j} + a_{i2} b_{2j} + \cdots + a_{in} b_{nj}$$

Linear Algebra

- Matrix Multiplication

○ **example:** computation of the coefficients c_{12} and c_{33} of the product matrix AB if A is (4,2)-dimension matrix, and B is a (2,3)-dimension matrix

○ $\forall i, j$: $c_{ij} = \sum_{k=1}^{n} a_{ik} b_{kj} = a_{i1} b_{1j} + a_{i2} b_{2j} + \cdots + a_{in} b_{nj}$



$$c_{12} = \sum_{k=1}^{n=2} a_{1k} b_{k2} = a_{11} b_{12} + a_{12} b_{22}$$

$$c_{33} = \sum_{k=1}^{n=2} a_{3k} b_{k3} = a_{31} b_{13} + a_{32} b_{23}$$

• Identity Matrix

- the **identity matrix** (or **unit matrix**) of size n is the $n \times n$ square matrix with ones on the main diagonal and zeros elsewhere
- since matrices can be multiplied on the condition that their types are compatible, there are unit matrices of any order
- I_n is a square matrix of order n is defined as a diagonal matrix with 1 of all entries of it main diagonal
- $I_1 = (1)$, $I_2 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$, $I_3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$, etc.
- $I_n A = A I_p = A$

Linear Algebra

- Matrix Transposition

- the transposed matrix of a matrix $A \in M_{(m,n)}(K)$ is the matrix denoted ${}^{t}A \in M_{(n,m)}(K)$ (also denoted A^t or A^{T}), obtained by reflecting A over its main diagonal (which runs from top-left to bottom-right) to obtain A^{T}
- if $B = A^{\mathrm{T}}$ then $\forall (i,j) \in \{1, \cdots, n\} \times \{1, \cdots, m\}$, $b_{j,i} = a_{i,j}$
- example : if $A = \begin{pmatrix} 1 & 3 & 5 \\ 2 & 4 & 6 \end{pmatrix}$ then $A^{\mathrm{T}} = \begin{pmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{pmatrix}$
- proprieties :
  - $\left(A^{\mathrm{T}}\right)^{\mathrm{T}} = A$
  - $(A + B)^{\mathrm{T}} = A^{\mathrm{T}} + B^{\mathrm{T}}$
  - for all scalar r, $(rA)^{\mathrm{T}} = rA^{\mathrm{T}}$
  - $(AB)^{\mathrm{T}} = B^{\mathrm{T}}A^{\mathrm{T}}$

Linear Algebra

• Matrix Determinant

- computation of the matrix determinant → useful to check if a matrix can be inverted or to compute the inverted matrix
- the general formula to compute the determinant
→ not easy for important size matrices
(→ but other techniques exist)
- for A, a (2×2) square matrix with $A = \begin{pmatrix} x & x' \\ y & y' \end{pmatrix}$,
the determinant is: $\det(A) = xy' - yx'$
- **geometric interpretation**: if $X = (x, y)$ and $X' = (x', y')$ are two vectors of Euclidian space, then the area of the parallelogram defined by X and X' is equal to the absolute value $xy' - yx'$ which is the determinant of the (2×2) square matrix $A = [X, X']$

- Matrix Determinant

  - the Leibniz formula expresses the determinant of a square matrix $A = \begin{pmatrix} a_{1;1} & \cdots & a_{1;n} \\ \vdots & \ddots & \vdots \\ a_{n;1} & \cdots & a_{n;n} \end{pmatrix}$ in terms of permutations of the matrix elements with the following formula:

$$\det(A) = |A| = \sum_{\sigma \in \wp_n} \mathcal{E}(\sigma) \prod_{i=1}^{n} a_{\sigma(i);i}$$

where $\wp_n$ denotes the permutations of $\{1, \cdots, n\}$ and $\mathcal{E}(\sigma)$ is the sign function of permutations σ in the permutation group $\wp_n$, which returns +1 and −1 for even and odd permutations, respectively

Linear Algebra

- Matrix Determinant

- for $n \geq 2$, the determinant of a square (n, n)-dimension matrix $A[a_{ij}]$ is the sum of n terms such as $\pm a_{ij} \times \det(A_{ij})$, with alternating plus or minus signs, with A_{ij} is the sub-matrix composed by removing the i^{th} row and j^{th} column of A :

$$\begin{aligned}\det(A) &= a_{11}\det(A_{11}) - a_{12}\det(A_{12}) + \cdots + (-1)^{1+n}a_{1n}\det(A_{1n}) \\ &= \sum_{j=1}^{n}(-1)^{1+j}a_{1j}\det(A_{1j})\end{aligned}$$

Linear Algebra

- Matrix Determinant

  - for a (3,3)-dimension square matrix $A\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}$
  - $\det(A) = a_{11}\det(A_{11}) - a_{12}\det(A_{12}) + a_{13}\det(A_{13})$
    $$\det(A) = a_{11}\begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix} - a_{12}\begin{vmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{vmatrix} + a_{13}\begin{vmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{vmatrix}$$
  - example: computation of the matrix $A\begin{pmatrix} 1 & 2 & 4 \\ 2 & -1 & 3 \\ 4 & 0 & 1 \end{pmatrix}$

$$\begin{vmatrix} 1 & 2 & 4 \\ 2 & -1 & 3 \\ 4 & 0 & 1 \end{vmatrix} = 1 \times \begin{vmatrix} -1 & 3 \\ 0 & 1 \end{vmatrix} - 2 \times \begin{vmatrix} 2 & 3 \\ 4 & 1 \end{vmatrix} + 4 \times \begin{vmatrix} 2 & -1 \\ 4 & 0 \end{vmatrix}$$
$$= 1 \times (-1) - 2 \times (-10) + 4 \times (4) = -1 + 20 + 16$$
$= 35 \neq 0$ therefore the matrix can be inverted

• Matrix Inversion

- an n-by-n square matrix A is called **invertible** (also "nonsingular" or "nondegenerate"), if there exists an n-by-n square B denoted $B = A^{-1}$ such that $AB = BA = I_n$ where I_n denotes the n-by-n identity matrix
- **theorem**: if $\det(A) \neq 0$, then A is invertible
- **theorem**: if $\det(A) = 0$, then A is not invertible, this is a **singular matrix**
- **properties**: if A and B are invertible matrices, then :

1. $(A^{-1})^{-1} = A$
2. $(AB)^{-1} = B^{-1}A^{-1}$
3. $\left(A^{\mathrm{T}}\right)^{-1} = (A^{-1})^{\mathrm{T}}$

- Matrix Inversion

- inverse of a 2×2 matrix:
- let $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$
- if $\det(A) \neq 0$, then the inverse matrix of A is:

$$A^{-1} = \frac{1}{\det(A)} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$$

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

Linear Algebra

- Matrix Inversion

○ example with the matrix inversion of a 3×3 square matrix:

○ let $A = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 2 & 3 \\ 1 & 2 & 4 \end{pmatrix}$

1. computation of the determinant of A :

$$\begin{aligned} \det(A) &= a_{11}\det(A_{11}) - a_{12}\det(A_{12}) + \cdots + (-1)^{1+n} a_{1n}\det(A_{1n}) \\ &= a_{11}\det(A_{11}) - a_{12}\det(A_{12}) + a_{13}\det(A_{13}) \\ &= 1 \times (8-6) - 1 \times (4-3) + 1 \times (2-2) = 1 \end{aligned}$$

2. computation of the transposed matrix A^{T} :

$$A^{\mathrm{T}} = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 2 & 2 \\ 1 & 3 & 4 \end{pmatrix}$$

Linear Algebra

- Matrix Inversion

○ example with a 3×3 square matrix (contined):

○ let $A = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 2 & 3 \\ 1 & 2 & 4 \end{pmatrix}$

1. computation of the determinant of A: $\det(A) = 1$
2. computation of the transposed matrix of A: $A^{\mathrm{T}} = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 2 & 2 \\ 1 & 3 & 4 \end{pmatrix}$
3. computation of the adjugate matrix A_{adj} from the determinants of each 2-by-2 matrices resulting from A^{T} :

$$A_{\mathrm{adj}} = \begin{pmatrix} 2 \times 4 - 3 \times 2 & 1 \times 4 - 1 \times 2 & 1 \times 3 - 1 \times 2 \\ 1 \times 4 - 3 \times 1 & 1 \times 4 - 1 \times 1 & 1 \times 3 - 1 \times 1 \\ 1 \times 2 - 2 \times 1 & 1 \times 2 - 1 \times 1 & 1 \times 2 - 1 \times 1 \end{pmatrix} = \begin{pmatrix} 2 & 2 & 1 \\ 1 & 3 & 2 \\ 0 & 1 & 1 \end{pmatrix}$$

Linear Algebra

- Matrix Inversion

- example with a 3×3 square matrix (continued):
- let $A = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 2 & 3 \\ 1 & 2 & 4 \end{pmatrix}$

4. sign matrix:

$$\begin{pmatrix} + & - & + \\ - & + & - \\ + & - & + \end{pmatrix}$$

5. computation of the inverse matrix A^{-1} by multiplying $\frac{1}{\det(A)} A_{\mathrm{adj}}$ by the sign matrix:

$$A^{-1} = \begin{pmatrix} 2 & -2 & 1 \\ -1 & 3 & -2 \\ 0 & -1 & 1 \end{pmatrix}$$

Linear Algebra

- Matrix Inversion

➢ **Algorithm for finding $\boldsymbol{A}^{-1}$, inverse matrix of $\boldsymbol{A}$**

- with elementary operations on the rows of A, transform A in the identity matrix I_n
- simultaneously perform the same operations on I_n
- at the end, A is transformed in I_n and I_n is transformed in A^{-1}
- in practice, we form the following table:

$$(A|I_n) \to (I_n|B) = (I_n|A^{-1})$$

- elementary operations authorized on the rows:

1. multiply a row by a non-zero real value: $L_i \leftarrow \lambda L_i$ with $\lambda \neq 0$
2. add to row L_i a multiple of L_j: $L_i \leftarrow L_i + \lambda L_j$
3. swap two rows: $L_i \leftrightarrows L_j$
4. warning: what is done on the left part of the augmented matrix must also be done on the right part

- Matrix Inversion

➢ **Algorithm for finding A^{-1}, inverse matrix of A** (example)

- compute the inverse matrix of $A = \begin{pmatrix} 1 & 2 & 1 \\ 4 & 0 & -1 \\ -1 & 2 & 2 \end{pmatrix}$
- here, the augmented matrix is:

$$(A|I_n) = \left(\begin{array}{ccc|ccc} 1 & 2 & 1 & 1 & 0 & 0 \\ 4 & 0 & -1 & 0 & 1 & 0 \\ -1 & 2 & 2 & 0 & 0 & 1 \end{array}\right) \begin{array}{l} L_1 \\ L_2 \\ L_3 \end{array}$$

$$\left(\begin{array}{ccc|ccc} 1 & 2 & 1 & 1 & 0 & 0 \\ 0 & -8 & -5 & -4 & 1 & 0 \\ -1 & 2 & 2 & 0 & 0 & 1 \end{array}\right) \begin{array}{c} L_1 \\ L_2 \leftarrow L_2 - 4L_1 \\ L_3 \end{array}$$

Linear Algebra

• Matrix Inversion

➢ **Algorithm for finding $\boldsymbol{A}^{-1}$, inverse matrix of $\boldsymbol{A}$** (continued)

$$\left(\begin{array}{ccc|ccc} 1 & 2 & 1 & 1 & 0 & 0 \\ 0 & -8 & -5 & -4 & 1 & 0 \\ 0 & 4 & 3 & 1 & 0 & 1 \end{array}\right) L_3 \leftarrow L_3 + L_1$$

$$\left(\begin{array}{ccc|ccc} 1 & 2 & 1 & 1 & 0 & 0 \\ 0 & 1 & \frac{5}{8} & \frac{1}{2} & -\frac{1}{8} & 0 \\ 0 & 4 & 3 & 1 & 0 & 1 \end{array}\right) L_2 \leftarrow -\frac{1}{8} L_2$$

$$\left(\begin{array}{ccc|ccc} 1 & 2 & 1 & 1 & 0 & 0 \\ 0 & 1 & \frac{5}{8} & \frac{1}{2} & -\frac{1}{8} & 0 \\ 0 & 0 & \frac{1}{2} & -1 & \frac{1}{2} & 1 \end{array}\right) L_3 \leftarrow L_3 - 4L_2$$

Linear Algebra

- Matrix Inversion

➢ **Algorithm for finding $\boldsymbol{A^{-1}}$, inverse matrix of $\boldsymbol{A}$** (continued)

$$\left(\begin{array}{ccc|ccc} 1 & 2 & 1 & 1 & 0 & 0 \\ 0 & 1 & \frac{5}{8} & \frac{1}{2} & -\frac{1}{8} & 0 \\ 0 & 0 & 1 & -2 & 1 & 2 \end{array}\right) L_3 \leftarrow 2L_3$$

$$\left(\begin{array}{ccc|ccc} 1 & 2 & 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & \frac{7}{4} & -\frac{3}{4} & -\frac{5}{4} \\ 0 & 0 & 1 & -2 & 1 & 2 \end{array}\right) L_2 \leftarrow L_2 - \frac{5}{8}L_3$$

$$\left(\begin{array}{ccc|ccc} 1 & 0 & 0 & -\frac{1}{2} & \frac{1}{2} & \frac{1}{2} \\ 0 & 1 & 0 & \frac{7}{4} & -\frac{3}{4} & -\frac{5}{4} \\ 0 & 0 & 1 & -2 & 1 & 2 \end{array}\right) L_1 \leftarrow L_1 - 2L_2 - L_3$$

Linear Algebra

- Matrix Inversion

➢ **Algorithm for finding $\boldsymbol{A^{-1}}$, inverse matrix of $\boldsymbol{A}$** (continued)

$$\left(\begin{array}{ccc|ccc} 1 & 0 & 0 & -\frac{1}{2} & \frac{1}{2} & \frac{1}{2} \\ 0 & 1 & 0 & \frac{7}{4} & -\frac{3}{4} & -\frac{5}{4} \\ 0 & 0 & 1 & -2 & 1 & 2 \end{array}\right)$$

therefore $A^{-1} = \frac{1}{4} \times \begin{pmatrix} -2 & 2 & 2 \\ 7 & -3 & -5 \\ -8 & 4 & 8 \end{pmatrix}$

we can check that $A \times A^{-1} = I_3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$

Linear Algebra

• Trace of a Matrix

- the **trace** of a $n \times n$ square matrix A, denoted $\mathrm{Tr}(A)$, is dened to be the sum of the elements on the main diagonal:

$$\mathrm{Tr}(A) = a_{11} + a_{22} + \cdots + a_{nn} = \sum_{i=1}^{n} a_{ii}$$

- the trace can be seen as a linear form on the vector space of the matrices
- the trace verifies the identity: $\mathrm{Tr}(AB) = \mathrm{Tr}(BA)$

- example :

➢ let a 2×2 square matrix $A = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}$

➢ $\rightarrow \mathrm{Tr}(A) = 2 + 2 = 4$

Linear Algebra

• Row Echelon Form of a Matrix

- a matrix is in echelon form if it has the shape resulting from a Gaussian elimination
- a matrix being in row echelon form means that Gaussian elimination has operated on the rows, and column echelon form means that Gaussian elimination has operated on the columns
- in other words, a matrix is in column echelon form if its transpose is in row echelon form
- the similar properties of column echelon form are easily deduced by transposing all the matrices

Linear Algebra

• Row Echelon Form of a Matrix

- specifically, a matrix is in row echelon form if
  - all rows consisting of only zeroes are at the bottom
  - the leading coefficient (also called the pivot) of a nonzero row is always strictly to the right of the leading coefficient of the row above it
- example where the $\divideontimes$ denote arbitrary coefficients and the $\oplus$ denotes the pivots (i.e., non-zero coefficients):

$$\begin{pmatrix} \oplus & \divideontimes & \divideontimes & \divideontimes & \divideontimes & \divideontimes & \divideontimes & \divideontimes & \divideontimes \\ 0 & 0 & \oplus & \divideontimes & \divideontimes & \divideontimes & \divideontimes & \divideontimes & \divideontimes \\ 0 & 0 & 0 & \oplus & \divideontimes & \divideontimes & \divideontimes & \divideontimes & \divideontimes \\ 0 & 0 & 0 & 0 & 0 & 0 & \oplus & \divideontimes & \divideontimes \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \oplus \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

Linear Algebra

- Row Echelon Form of a Matrix

➢ **Transformation to row echelon form**

- by means of a finite sequence of elementary row operations, any matrix can be transformed to row echelon form
- since elementary row operations preserve the row space of the matrix, the row space of the row echelon form is the same as that of the original matrix
- authorized operations (→ Gaussian elimination):

1. swap two lines
2. multiply a row by a non-zero constant
3. add to a row the multiple of another row

- the number of rows with a non-zero pivot is equal to the **rank** of the initial matrix (→ rank = maximal number of linearly independent columns of the matrix)

Linear Algebra

• Vector space, image and kernel

- let A be a matrix with m rows and n columns, and rank r (the rank corresponds to the dimension of the vector space spanned by its rows)
- let C be the matrix constituted by the r first rows at the row echelon form of the corresponding matrix (the following rows are equal to zero)
- the transformation to row echelon form is made by blocs $(A|I_m)$ where I_m is the identity matrix with m rows
- let $\begin{pmatrix} C & K \\ 0 & L \end{pmatrix}$ the corresponding bloc row echelon form matrix
- the matrices C and L allow to determine some sub-spaces associated to the matrix A
- in the case of $m = n = r$, the matrix K is A^{-1}

Linear Algebra

- Vector space, image and kernel

➢ Kernel of a matrix

- the **kernel** of a linear mapping, also known as the null space, is the set of vectors in the domain of the mapping which are mapped to the zero vector
- the **kernel** $\mathrm{Ker}(A)$ of matrix A is defined as the vector subspace of $\mathbb{K}^n$ constituted by the columns X solutions of the linear system $AX = 0$

- if $(e_1, e_2, \cdots, e_n)$ are the components of the basis of $\mathbb{K}^n$ and $(k_1, k_2, \cdots, k_n)$ are the indices of the columns with the pivots,
- then a basis of $\mathrm{Ker}(A)$ is given by $e_j - \sum_{k_i<j} c_{ij} e_{ki}$

- Vector space, image and kernel

➢ **Image of a matrix**

○ the **image** $\mathrm{Im}(A)$ of matrix A :

❑ vector subspace of $\mathbb{K}^m$

❑ constituted by AX

❑ when X is a column with n terms

○ this image is generated by the columns of A, and a basis is formed by the columns whose index contains, after reduction, a pivot (i.e., a non-zero coefficient)

Linear Algebra

• Vector space, image and kernel

- example with a rectangular matrix characterizing a linear map f:

$$\begin{pmatrix} 1 & 3 & 1 \\ -1 & 0 & 2 \\ 2 & 5 & 1 \\ 1 & 3 & 1 \end{pmatrix}$$

- here, the linear map is $f\colon \mathbb{R}^3 \to \mathbb{R}^4$
- $$\begin{pmatrix} 1 & 3 & 1 \\ -1 & 0 & 2 \\ 2 & 5 & 1 \\ 1 & 3 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} x + 3y + z \\ -x + 2z \\ 2x + 5y + z \\ x + 3y + z \end{pmatrix}$$

Linear Algebra

- Vector space, image and kernel

- for the kernel, we look for x, y and z which verify the following equations:

$$\begin{pmatrix} x & +3y & +z = 0 \\ -x & & +2z = 0 \\ 2x & +5y & +z = 0 \\ x & +3y & +z = 0 \end{pmatrix}$$

- for the image, we are looking for a relation between x', y', z' and t':

$$\begin{pmatrix} x & +3y & +z = x' \\ -x & & +2z = y' \\ 2x & +5y & +z = z' \\ x & +3y & +z = t' \end{pmatrix}$$

Linear Algebra

- Eigenvectors and Eigenvalues

- the **eigenvectors** of a square matrix A are the vectors other than zero which, after multiplication by A, remain parallel to the original vector
- for any eigenvector $\vec{x}$, there exists a corresponding **eigenvalue** λ which is the factor by which the eigenvector $\vec{x}$ is resized by multiplication with A
- $\vec{x}$ is an eigenvector of A if there is a scalar λ such that $Ax = \lambda x$
- with T, transformation by linear map:

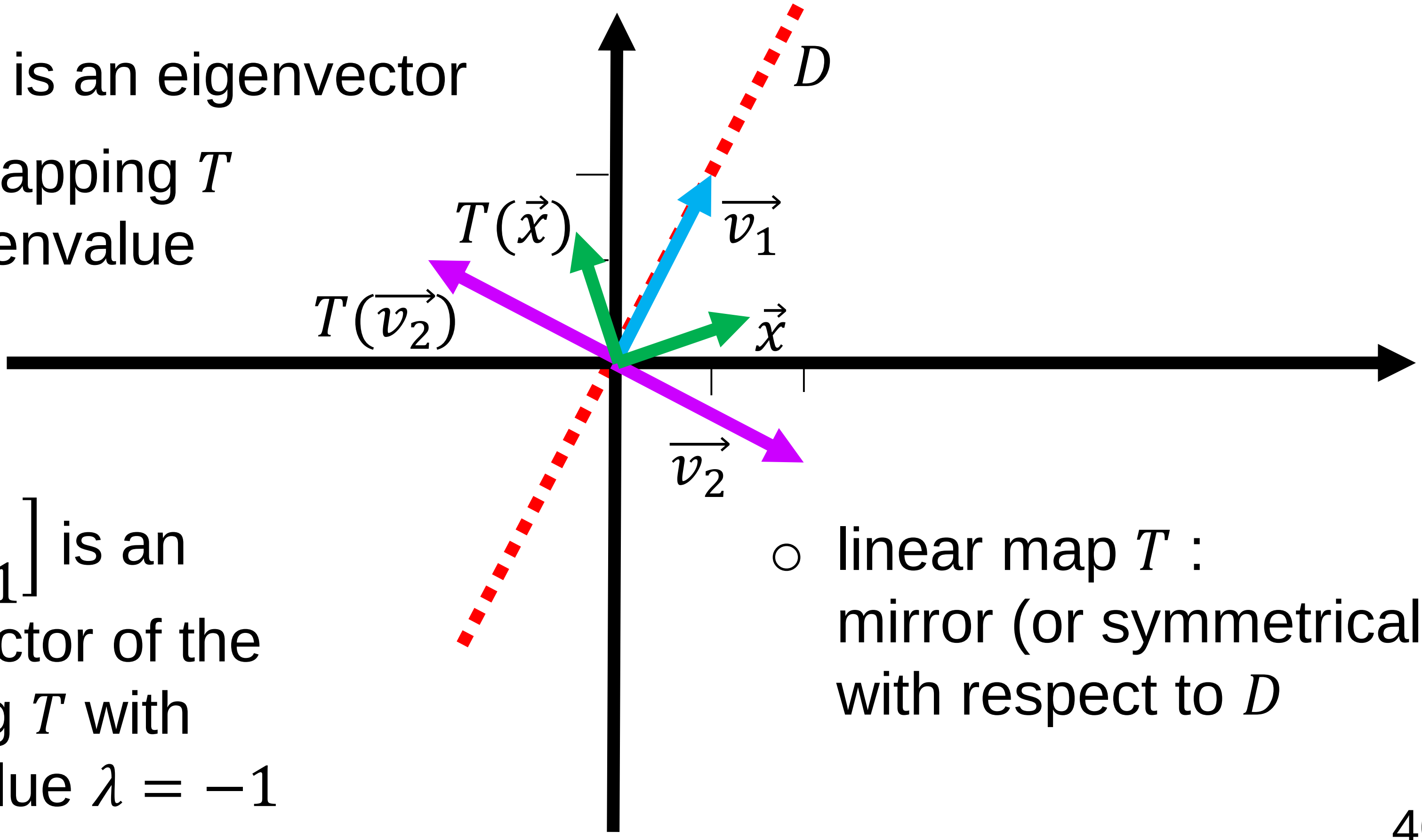
$$T: \mathbb{R}^n \rightarrow \mathbb{R}^n$$
$$T(\vec{v}) = \lambda\vec{v}$$

- the vector $\vec{v}$, when we apply the transformation T to it, is only modified by a factor λ (it is just made smaller or larger)

Linear Algebra

- Eigenvectors and Eigenvalues

- example: linear map $T: \mathbb{R}^2 \rightarrow \mathbb{R}^2$
- $\overrightarrow{v_1} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$ is an eigenvector of the mapping T with eigenvalue $\lambda = 1$



- $\overrightarrow{v_2} = \begin{bmatrix} 2 \\ -1 \end{bmatrix}$ is an eigenvector of the
- mapping T with eigenvalue $\lambda = -1$
- linear map T : mirror (or symmetrical) with respect to D

Linear Algebra

- Eigenvectors and Eigenvalues

  - $T(\overrightarrow{v_1}) = \overrightarrow{v_1} \Longrightarrow T(\overrightarrow{v_1}) = 1 \times \overrightarrow{v_1}$
  - $T(\overrightarrow{v_2}) = -\overrightarrow{v_2} \Longrightarrow T(\overrightarrow{v_2}) = -1 \times \overrightarrow{v_2}$
  - by mapping T, the vectors $\overrightarrow{v_1}$ and $\overrightarrow{v_2}$ will not change a lot: the orientation will be the same but the direction or the amplitude can change
  - the vectors $\overrightarrow{v_1}$ and $\overrightarrow{v_2}$ are the eigenvectors of the mapping T and λ is the corresponding eigenvalue ($\lambda_1 = 1$ and $\lambda_2 = -1$)
  - **interest**: in a basis defined by two vectors $\overrightarrow{v_1}$ and $\overrightarrow{v_2}$, the corresponding matrix to the linear map T can easily been expressed
  - eigenvectors are good candidates for a basis of space and for simply expressing the linear map

Linear Algebra

• Eigenvectors and Eigenvalues

- in the general case, with a matrix A :

$$T(\vec{x}) = A\vec{x}$$
$$T(\vec{v}) = \lambda\vec{v} = A\vec{v}$$

- we will say that $\vec{v}$ is an eigenvector associated with the linear map T but $\vec{v}$ is also an eigenvector of A, and the scalar λ is an eigenvalue of A

- therefore we have:

$$A\vec{v} = \lambda\vec{v}$$

- what are the solutions of this equation?
- we will not be interested in the obvious solution $\vec{v} = \vec{0}$

- Eigenvectors and Eigenvalues

- in the general case, with a matrix A, the eigenvalues of A are the solutions λ of the characteristic equation:

$$\det(A - \lambda I_n) = 0$$

- other writing:
  $A\vec{v} = \lambda\vec{v}$ for $\vec{v} \neq \vec{0}$
  iff $\det(\lambda I_n - A) = 0$

- λ is an eigenvalue of A
  iff $\det(\lambda I_n - A) = 0$

Linear Algebra

- Eigenvectors and Eigenvalues

- example with a 2 × 2 square matrix:
- $A = \begin{bmatrix} 1 & 2 \\ 4 & 3 \end{bmatrix}$ and suppose that λ is an eigenvalue of A
- $\det(\lambda I_2 - A) = 0$

$$\Leftrightarrow \det\left(\lambda \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} - \begin{bmatrix} 1 & 2 \\ 4 & 3 \end{bmatrix}\right) = 0$$

$$\Leftrightarrow \det\left(\begin{bmatrix} \lambda & 0 \\ 0 & \lambda \end{bmatrix} - \begin{bmatrix} 1 & 2 \\ 4 & 3 \end{bmatrix}\right) = 0$$

$$\Leftrightarrow \det\left(\begin{bmatrix} \lambda - 1 & -2 \\ -4 & \lambda - 3 \end{bmatrix}\right) = 0 \quad \text{with } \det\begin{pmatrix} x & x' \\ y & y' \end{pmatrix} = xy' - yx'$$

$$\Leftrightarrow (\lambda - 1)(\lambda - 3) - (-4 \times -2) = 0$$

$$\Leftrightarrow (\lambda - 1)(\lambda - 3) - 8 = 0$$

$$\Leftrightarrow \lambda^2 - 3\lambda - \lambda + 3 - 8 = 0$$

$$\Leftrightarrow \lambda^2 - 4\lambda - 5 = 0$$

Linear Algebra

• Eigenvectors and Eigenvalues

- example with a 2 × 2 square matrix $A = \begin{bmatrix} 1 & 2 \\ 4 & 3 \end{bmatrix}$ (continued):
- $\lambda^2 - 4\lambda - 5 = 0$ → we call this polynomial the "characteristic polynomial" of A
- we have to find two value λ_1 and λ_2 such that their product will be -5 and their sum will be 4
- → solutions of a quadratic equation
- $\lambda^2 - 4\lambda - 5 = 0 \rightarrow \Delta = b^2 - 4ac = (-4)^2 - 4\big(1 \times (-5)\big)$
- $\Delta = 16 + 20 = 36 = 6^2$
- solutions:

➢ $\lambda_1 = \frac{-b+\sqrt{\Delta}}{2a} = \frac{+4+6}{2\times 1} = \frac{10}{2} = 5$

➢ $\lambda_2 = \frac{-b-\sqrt{\Delta}}{2a} = \frac{+4-6}{2\times 1} = \frac{-2}{2} = -1$

Linear Algebra

- Eigenvectors and Eigenvalues

- example with a 2 × 2 square matrix $A = \begin{bmatrix} 1 & 2 \\ 4 & 3 \end{bmatrix}$ (continued):
- $\lambda^2 - 4\lambda - 5 = 0$ → eigenvalues: $\lambda_1 = 5$ and $\lambda_2 = -1$

- to find the eigenvectors associated with the eigenvalues, we start from the basic equation $A\vec{v} = \lambda\vec{v}$ with $\vec{v} \neq \vec{0}$

$$\Leftrightarrow \vec{0} = \lambda\vec{v} - A\vec{v}$$
$$\Leftrightarrow \vec{0} = \lambda I_n\vec{v} - A\vec{v}$$
$$\Leftrightarrow \vec{0} = \lambda(I_n - A)\vec{v}$$

- for any eigenvalue λ, the eigenspace associated with the eigenvalue λ is $E_\lambda = \mathrm{Ker}(\lambda I_n - A)$

Linear Algebra

- Eigenvectors and Eigenvalues

- example with a 2 × 2 square matrix $A = \begin{bmatrix} 1 & 2 \\ 4 & 3 \end{bmatrix}$ (continued):
- with $E_\lambda = \mathrm{Ker}(\lambda I_n - A)$, pour $\lambda_1 = 5$:
- $E_5 = \mathrm{Ker}(5\, I_2 - A) = \mathrm{Ker}\left(\begin{bmatrix} 5 & 0 \\ 0 & 5 \end{bmatrix} - \begin{bmatrix} 1 & 2 \\ 4 & 3 \end{bmatrix}\right) = \mathrm{Ker}\left(\begin{bmatrix} 4 & -2 \\ -4 & 2 \end{bmatrix}\right)$
- we try to find the kernel of this matrix
- → we compute the reduced row echelon form of this matrix
- $\begin{bmatrix} 4 & -2 \\ -4 & 2 \end{bmatrix} \begin{array}{l} L_1 \\ L_2 \end{array} \longrightarrow \begin{bmatrix} 4 & -2 \\ 0 & 0 \end{bmatrix}_{L_2 \leftarrow L_2 + L_1} \longrightarrow \begin{bmatrix} 1 & -\frac{1}{2} \\ 0 & 0 \end{bmatrix}^{L_1 \leftarrow L_1/4}$
- we multiply the reduced row echelon form matrix by a vector $\begin{bmatrix} v_1 \\ v_2 \end{bmatrix}$ and we establish that the matrix product is zero
- $\begin{bmatrix} 1 & -\frac{1}{2} \\ 0 & 0 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$

Linear Algebra

- Eigenvectors and Eigenvalues

- example with a 2 × 2 square matrix $A = \begin{bmatrix} 1 & 2 \\ 4 & 3 \end{bmatrix}$ (continued):
- with $E_\lambda = \mathrm{Ker}(\lambda I_n - A)$, for $\lambda_1 = 5$:
- $\begin{bmatrix} 1 & -\frac{1}{2} \\ 0 & 0 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \Leftrightarrow v_1 - \frac{1}{2} v_2 = 0 \Leftrightarrow v_1 = \frac{1}{2} v_2$
- if we say that $v_2 = t$ then $v_1 = \frac{1}{2} t$
- we can rewrite the eigenspace associated with the eigenvalue 5 which is the kernel of this matrix:

$$E_5 = \left\{ \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} = t \begin{bmatrix} \frac{1}{2} \\ 1 \end{bmatrix}, t \in \mathbb{R} \right\}$$

$$E_5 = \mathrm{Vect}\left(\begin{bmatrix} \frac{1}{2} \\ 1 \end{bmatrix} \right)$$

Linear Algebra

- Eigenvectors and Eigenvalues

- example with a 2×2 square matrix $A = \begin{bmatrix} 1 & 2 \\ 4 & 3 \end{bmatrix}$ (continued):
- with $E_\lambda = \mathrm{Ker}(\lambda I_n - A)$, for $\lambda_2 = -1$:
- $E_{-1} = \mathrm{Ker}(-I_2 - A) = \mathrm{Ker}\left(\begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix} - \begin{bmatrix} 1 & 2 \\ 4 & 3 \end{bmatrix}\right) =$
  $\mathrm{Ker}\left(\begin{bmatrix} -2 & -2 \\ -4 & -4 \end{bmatrix}\right)$
- we transform the matrix in its reduced row echelon form
- $\begin{bmatrix} -2 & -2 \\ -4 & -4 \end{bmatrix} \begin{matrix} L_1 \\ L_2 \end{matrix} \longrightarrow \begin{bmatrix} -2 & -2 \\ 0 & 0 \end{bmatrix}_{L_2 \leftarrow L_2 - 2L_1} \longrightarrow \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}^{L_1 \leftarrow -L_1/4}$
- to have the kernel, we multiply the matrix by the vector $\begin{bmatrix} v_1 \\ v_2 \end{bmatrix}$
- $\begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$

- Eigenvectors and Eigenvalues

○ example with a 2 × 2 square matrix $A = \begin{bmatrix} 1 & 2 \\ 4 & 3 \end{bmatrix}$ (continued):

○ with $E_\lambda = \mathrm{Ker}(\lambda I_n - A)$, for $\lambda_2 = -1$:

○ $\begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \Leftrightarrow v_1 + v_2 = 0 \Leftrightarrow v_1 = -v_2$

○ if we say that $v_2 = t$ then $v_1 = -t$

○ we can rewrite the eigenspace associated with the eigenvalue -1 which is the kernel of this matrix:

$$E_{-1} = \left\{ \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} = t \begin{bmatrix} -1 \\ 1 \end{bmatrix}, t \in \mathbb{R} \right\}$$

$$E_{-1} = \mathrm{Vect}\left(\begin{bmatrix} -1 \\ 1 \end{bmatrix} \right)$$

Linear Algebra

- Eigenvectors and Eigenvalues

- example with a 3×3 square matrix:
- $A = \begin{bmatrix} -1 & 2 & 2 \\ 2 & 2 & -1 \\ 2 & -1 & 2 \end{bmatrix}$ and λ is an eigenvalue of A
- $\det(\lambda I_3 - A) = 0$

$$\Leftrightarrow \det\left(\lambda \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} - \begin{bmatrix} -1 & 2 & 2 \\ 2 & 2 & -1 \\ 2 & -1 & 2 \end{bmatrix}\right) = 0$$

$$\Leftrightarrow \det\left(\begin{bmatrix} \lambda + 1 & -2 & -2 \\ -2 & \lambda - 2 & 1 \\ -2 & 1 & \lambda - 2 \end{bmatrix}\right) = 0$$

- to get the determinant of a 3×3 matrix is less easy than a 2×2 matrix

Linear Algebra

- Eigenvectors and Eigenvalues

○ example with a 3 × 3 square matrix (continued):

○ method to obtain the determinant: consider the products of the coefficients on the diagonals by alternating the sign in the sum of the products and repeating the start of the matrix

○ $$\left[\begin{array}{ccc|cc} \lambda+1 & -2 & -2 & \lambda+1 & -2 \\ -2 & \lambda-2 & 1 & -2 & \lambda-2 \\ -2 & 1 & \lambda-2 & -2 & 1 \end{array}\right]$$

$$\Rightarrow (\lambda+1)(\lambda-2)(\lambda-2)$$
$$\Rightarrow (-2)\times 1\times(-2) = 4$$
$$\Rightarrow (-2)\times(-2)\times 1 = 4$$

○ $$\left[\begin{array}{ccc|cc} \lambda+1 & -2 & -2 & \lambda+1 & -2 \\ -2 & \lambda-2 & 1 & -2 & \lambda-2 \\ -2 & 1 & \lambda-2 & -2 & 1 \end{array}\right]$$

$$\Rightarrow (-2)\times(\lambda-2)\times(-2) = 4(\lambda-2)$$
$$\Rightarrow 1\times 1\times(\lambda+1) = \lambda+1$$
$$\Rightarrow (\lambda-2)\times(-2)\times(-2) = 4(\lambda-2)$$

$$(\lambda+1)(\lambda-2)(\lambda-2)+4+4-4(\lambda-2)-(\lambda+1)-4(\lambda-2) = 0$$

- Eigenvectors and Eigenvalues

- example with a 3 × 3 square matrix (continued):

$$(\lambda+1)(\lambda-2)(\lambda-2)+4+4-4(\lambda-2)-(\lambda+1)-4(\lambda-2)=0$$
$$\Leftrightarrow (\lambda+1)(\lambda^2-4\lambda+4)+8-4\lambda+8-\lambda-1-4\lambda+8=0$$
$$\Leftrightarrow \lambda^3-4\lambda^2+4\lambda+\lambda^2-4\lambda+4-9\lambda+23=0$$
$$\Leftrightarrow \lambda^3-3\lambda^2-9\lambda+27=0$$

- this is the characteristic polynomial of the matrix A
- this equation is not easy to solve, but the roots of this equation are bound to be divisors of 27, i.e., necessarily 1, 3, 9 or 27
- we will test these roots to see if they work:
- 1: 1−3−9 + 27 ≠ 0 so 1 is not root
- 3: 27−3 × 9−9 × 3 + 27 = 0 so 3 is a root
- we can factorize by $(\lambda - 3)$

Linear Algebra

- Eigenvectors and Eigenvalues

- example with a 3 × 3 square matrix (continued):
- $\lambda^3 - 3\lambda^2 - 9\lambda + 27 = 0$ and we can factorize by $(\lambda - 3)$
- $(\lambda - 3)(\lambda^2 - 9) = 0 \Leftrightarrow (\lambda - 3)(\lambda - 3)(\lambda + 3)$
- so we get the eigenvalues $\lambda = 3$ or $\lambda = -3$
- we try to find now the eigenvectors and the eigenspaces associated with the eigenvalues
- the eigenvectors belong to the kernel of the matrix A
- for $\lambda = 3$, $\lambda I_n - A$:
- $3\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} - \begin{bmatrix} -1 & 2 & 2 \\ 2 & 2 & -1 \\ 2 & -1 & 2 \end{bmatrix} = \begin{bmatrix} 4 & -2 & -2 \\ -2 & 1 & 1 \\ -2 & 1 & 1 \end{bmatrix}$

Linear Algebra

- Eigenvectors and Eigenvalues

- example with a 3 × 3 square matrix (continued):
- for $\lambda = 3$, $\lambda I_n - A$:
- $$\begin{bmatrix} 4 & -2 & -2 \\ -2 & 1 & 1 \\ -2 & 1 & 1 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$
- we put the matrix in reduced row echelon form to be able to easily find the kernel of this matrix
- we keep the 1st row and we will replace the 2nd row by 2 times the 2nd row + the 1st row (the same for the 3rd row)
- then we divide the 1st row by 4
- $$\begin{bmatrix} 1 & -\frac{1}{2} & -\frac{1}{2} \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

Linear Algebra

- Eigenvectors and Eigenvalues

  - example with a 3 × 3 square matrix (continued):
  - $\begin{bmatrix} 1 & -\frac{1}{2} & -\frac{1}{2} \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$ therefore $v_1 - \frac{1}{2} v_2 - \frac{1}{2} v_3 = 0$
  - if $v_2 = a$, $v_3 = b$ then $v_1 = \frac{1}{2} a + \frac{1}{2} b$
  - → eigenspace corresponding to the kernel of the matrix
  - $E_3 = \left\{ \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} = a \begin{bmatrix} \frac{1}{2} \\ 1 \\ 0 \end{bmatrix} + b \begin{bmatrix} \frac{1}{2} \\ 0 \\ 1 \end{bmatrix}, a, b \in \mathbb{R} \right\}$
  - $E_3 = \mathrm{Vect}\left(\begin{bmatrix} \frac{1}{2} \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} \frac{1}{2} \\ 0 \\ 1 \end{bmatrix} \right)$

Linear Algebra

- Eigenvectors and Eigenvalues

○ example with a 3 × 3 square matrix (continued):

○ now we try to find the eigenspace associated to the eigenvalue $\lambda = -3$, $\lambda I_n - A$:

○ $-3\begin{bmatrix}1 & 0 & 0\\0 & 1 & 0\\0 & 0 & 1\end{bmatrix} - \begin{bmatrix}-1 & 2 & 2\\2 & 2 & -1\\2 & -1 & 2\end{bmatrix} = \begin{bmatrix}-2 & -2 & -2\\-2 & -5 & 1\\-2 & 1 & -5\end{bmatrix}$

○ we divide the 1st row by −2 then we replace the 2nd row by the 2nd row − the 1st row (the same for the 3rd row)

○ we divide the 2nd row by −3 then we replace the 3rd row by the 3rd row + the 2nd row

○ $\begin{bmatrix}1 & 1 & 1\\0 & 1 & -1\\0 & 0 & 0\end{bmatrix}\begin{bmatrix}v_1\\v_2\\v_3\end{bmatrix} = \begin{bmatrix}0\\0\\0\end{bmatrix}$

Linear Algebra

- Eigenvectors and Eigenvalues

- example with a 3 × 3 square matrix (continued):
- $\begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & -1 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$ therefore $\begin{cases} v_1 + v_2 + v_3 = 0 \\ v_2 - v_3 = 0 \end{cases}$
- we say that v_3 is the free variable : $v_3 = t$
- then $v_2 = t$ et $v_1 = -2t$
- → eigenspace corresponding to the kernel of the matrix
- $E_{-3} = \left\{ \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} = t \begin{bmatrix} -2 \\ 1 \\ 1 \end{bmatrix}, t \in \mathbb{R} \right\}$
- $E_{-3} = \mathrm{Vect}\left(\begin{bmatrix} -2 \\ 1 \\ 1 \end{bmatrix} \right)$

- Eigenvectors and Eigenvalues

  - example with a 3×3 square matrix (continued):
  - remarkable element: if we make the scalar product of $\begin{bmatrix} -2 \\ 1 \\ 1 \end{bmatrix}$ of the eigenspace E_{-3} with one of the eigenvectors of E_3, i.e., $\begin{bmatrix} \frac{1}{2} \\ 1 \\ 0 \end{bmatrix}$ or $\begin{bmatrix} \frac{1}{2} \\ 0 \\ 1 \end{bmatrix}$, we get 0
  - therefore the vectors are orthogonal
  - → the eigenspace associated with the eigenvalue -3 is orthogonal to the eigenspace associated with the value 3

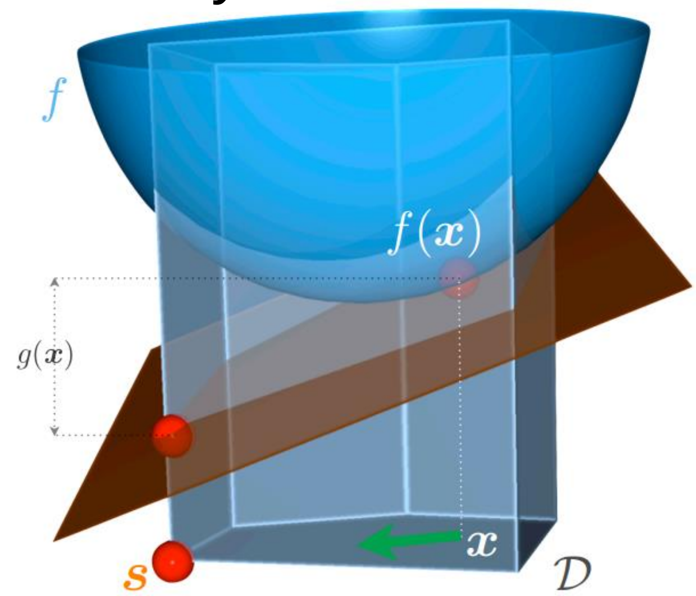
Convex Optimization

- Introduction

➢ **Convex optimization →**
problem of minimizing convex functions over convex sets

➢ the convexity property can make optimization in some sense "easier" than the general case: indeed, any local minimum must be a global minimum.

➢ **Objective:** find a stationary point



Convex Optimization

- Convex optimization and linear algebra

➢ **convex optimization** problem:

$$\begin{cases} max_{\mathbf{u}} \mathbf{u}^T S \mathbf{u} & (1) \\ \text{subject to } \mathbf{u}^T \mathbf{u} = 1 & (2) \end{cases}$$

➢ Solving problem (1) under the constraint (2), we get: $S\mathbf{u} = \lambda\mathbf{u}$, which says that $\mathbf{u}$ must be an **eigenvector** of S.

➢ We can deduce that $\mathbf{u}^T S \mathbf{u} = \lambda$,
and so the variance will be maximum when we set $\mathbf{u}$ equal to the eigenvector having the largest **eigenvalue** λ.

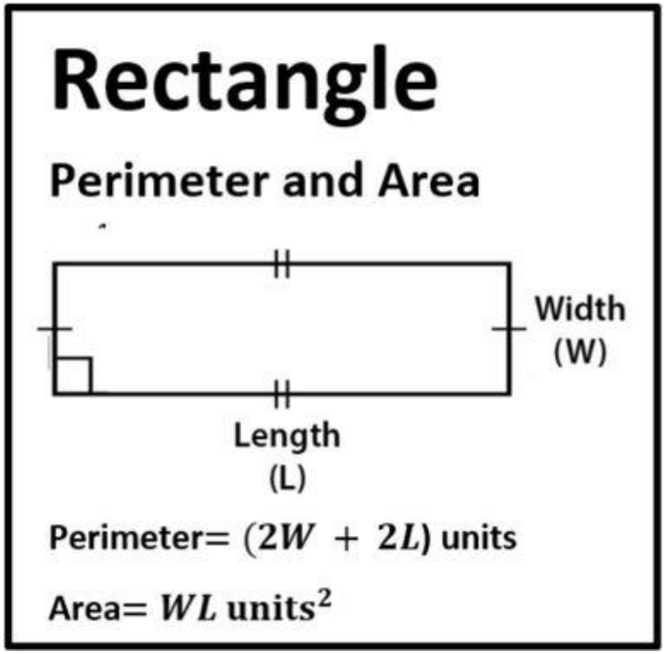
➢ Note that the eigenvalues of a matrix S are the solutions to the following characteristic equation:
$det(S - \lambda I) = 0$, where det is the **determinant**.

Convex Optimization

- Introductory example: length & width of a rectangle

- let *R* be a rectangle of width *W* and length *L*
- let us solve the following problem: “Find *W* and *L* which minimize the perimeter *P* of *R* under the constraint that the area *A* of *R* is equal to 4”
- more formally, this takes the form of the following optimization problem:

$$\begin{cases} min_{W,L} 2W + 2L \\ R = W \times L = 4 \end{cases}$$



Convex Optimization

- Introductory example: length & width of a rectangle

- the method of **Lagrange multipliers** tells us that solving the previous optimization problem boils down to minimizing the following problem: $min_{W,L} 2W + 2L + \lambda(W \times L - 4)$ for some λ
- solving $\nabla\big(2W + 2L + \lambda(W \times L - 4)\big) = 0$
  we get:

$$\begin{cases} \dfrac{\partial\big(2W+2L+\lambda(W\times L-4)\big)}{\partial W} = 0 \Leftrightarrow 2 + \lambda W = 0 \Leftrightarrow W = -\dfrac{2}{\lambda} & (1) \\ \dfrac{\partial\big(2W+2L+\lambda(W\times L-4)\big)}{\partial L} = 0 \Leftrightarrow 2 + \lambda L = 0 \Leftrightarrow L = -\dfrac{2}{\lambda} & (2) \\ \dfrac{\partial\big(2W+2L+\lambda(W\times L-4)\big)}{\partial \lambda} = 0 \Leftrightarrow W \times L - 4 = 0 \Leftrightarrow W \times L = 4 & (3) \end{cases}$$

Convex Optimization

- Introductory example: length & width of a rectangle

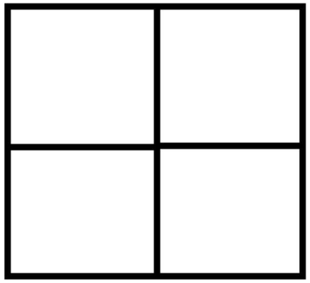
  - from $\begin{cases} W = -\frac{2}{\lambda} & (1) \\ L = -\frac{2}{\lambda} & (2) \\ W \times L = 4 & (3) \end{cases}$

    plugging Eq.(1) and (2) into Eq.(3), we obtain $\frac{4}{\lambda^2} = 4$

    therefore $\lambda^2 = 1$, we get that $\lambda = 1$ or $\lambda = -1$

    and we deduce that $Width = 2$ and $Length = 2$

  - → perimeter $P = 2W + 2L = 8$



Convex Optimization

- Definition

➢ Optimization Problem → determine value of **optimization variable** within **feasible region/set** to optimize **optimization objective**

$$\min_{x} f(x)$$
$$\text{s. t. } \; x \in \mathcal{F}$$

- optimization variable $x \in \mathbb{R}^n$ (x: vector in n real-valued entries)
- feasible region/set $\mathcal{F} \subset \mathbb{R}^n$
- optimization objective $f: \mathcal{F} \to \mathbb{R}$

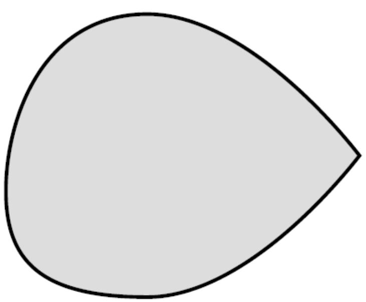
- Optimal solution: $x^* = \operatorname{argmin}_{x \in \mathcal{F}} f(x)$
- Optimal objective value $f^* = \min_{x \in \mathcal{F}} f(x) = f(x^*)$

Convex Optimization

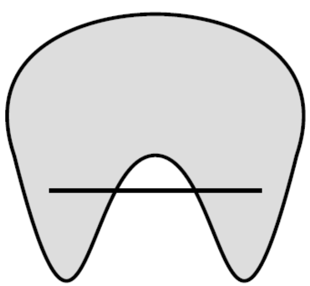
- Definition

$$\min_{x} f(x)$$
$$\text{s. t. } \; x \in \mathcal{F}$$

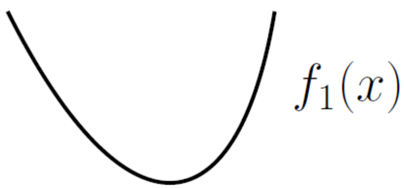
- An optimization problem whose optimization objective f is a **convex function** and feasible region $\mathcal{F}$ is a **convex set**
- → a special class of optimization problem



Convex set



Nonconvex set



Convex function

Nonconvex function

Convex Optimization

- Definition: Local optima and global optima

$$\min_x f(x)$$
$$\text{s. t. } \; x \in \mathcal{F}$$

- ➢ Given an optimization problem, a point $x \in \mathbb{R}^n$ is **globally optimal** if $x \in \mathcal{F}$ and $\forall y \in \mathcal{F}, f(x) \leq f(y)$
- ➢ Given an optimization problem, a point $x \in \mathbb{R}^n$ is **locally optimal** if $x \in \mathcal{F}$ and $\exists R > 0$ such that $\forall y: y \in \mathcal{F}$ and $\|x - y\|_2 \leq R, f(x) \leq f(y)$

- ➢ **Theorem**: for a convex optimization problem, all locally optimal points are globally optimal

Convex Optimization

- Definition

$$\min_{x} f(x)$$
$$\text{s. t.} \quad x \in \mathcal{F}$$

➢ Optimization variable $x \in \mathbb{R}^n$

- discrete variables → combinatorial optimization
- continuous variables → Continuous optimization
- mixed → some variables are discrete, and some continuous

- example: shortest path, traveling salesman problem...

Convex Optimization

- Definition

$$\min_{x} f(x)$$
$$\text{s. t. } \; x \in \mathcal{F}$$

- ➢ Feasible region/set $\mathcal{F} \subset \mathbb{R}^n$
- ○ unconstrained optimization: $\mathcal{F} = \mathbb{R}^n$
- ○ constrained optimization: $\mathcal{F} \subsetneq \mathbb{R}^n$

  → find a feasible point $x \in \mathcal{F}$ can already be difficult

Convex Optimization

- Definition

$$\min_x f(x)$$
$$\text{s. t. } \; x \in \mathcal{F}$$

➢ Optimization objective $f: \mathcal{F} \to \mathbb{R}$

○ $f(x) = 1$: feasibility problem

○ simple functions:

❑ linear function $f(x) = a^T x$

❑ convex function

○ complicated functions

❑ even can be implicitly represented through an algorithm which takes $x \in \mathcal{F}$ as input, and outputs a value

Convex Optimization

- Definition

$$\min_{x} f(x)$$
$$\text{s. t.} \quad x \in \mathcal{F}$$

➢ Minimization can be converted to maximization (and vice versa): $\max_{x} g(x) = -f(x)$

s.t. $x \in \mathcal{F}$

○ same optimal solution
optimal objective value $g^* = -f^*$

- Example 1: Traveling Salesman Problem (TSP)

$$\min_x f(x)$$
$$\text{s. t. } \; x \in \mathcal{F}$$

- Problem: n cities, distance from city i to city j is $d(i,j)$, find a tour (a closed path that visits every city exactly once) with minimal total distance
- Variable x: ordered list of cities being visited
  - x_i is the index of the i^{th} city being visited
- Feasible set F={x: each city visited exactly once}
  - $F = \{x\colon x \in \{1..n\}^n; \sum_k \mathbb{1}(x_k = i) = 1, \forall i \in \{1..n\}\}$
- Objective function $f(x)$= total distance when following x
  - $f(x) = \sum_{k=1}^{n-1} d(x_k, x_{k+1}) + d(x_n, x_1)$

Convex Optimization

- Example 2: 8-Queens Problem

$$\min_{x} f(x)$$
$$\text{s. t. } \; x \in \mathcal{F}$$

- Problem: placing eight chess queens on an 8×8 chessboard so that no two queens threaten each other
(→ two queens don't share the same row/column/diagonal)
- Variable x: location of the queen in each column
  - x_i is the row index of the queen in i^{th} column
- Feasible set F={x: no queens in the row, col, diag}
  - $F = \{x, y : x, y \in \{1..8\}^8;\ \sum_i \mathbb{I}(x_i = k) = 1, \forall k \in \{1..8\};$
$\sum_i \mathbb{I}(y_i = k) = 1, \forall k \in \{1..8\}; |x_i - x_j| \neq |y_i - y_j|, \forall i, j \in \{1..8\}\}$
- Objective function $f(x) = 1$ (dummy)

Convex Optimization

• Example 3: Linear Regression

$$\min_a f(a)$$
$$\text{s. t. } \; a \in \mathbb{R}$$

x_i	1.0	2.0	3.5
y_i	2.1	3.98	7.0

- Problem: Find a such that $y_i \approx ax_i, \forall i = 1..3$
- Variable a
- Feasible region $\mathbb{R}$
- Objective function $f(a)$?

$$\min_a \sum_{i=1}^{3} |y_i - ax_i| \qquad \text{s.t. } a \in \mathbb{R}$$

$$\min_a \sum_{i=1}^{3} (y_i - ax_i)^2 \qquad \text{s.t. } a \in \mathbb{R}$$

Convex Optimization

- How to determine if a function is convex

➢ Prove by definition

➢ Use properties:

- Sum of convex functions is convex
  - If $f(x) = \sum_i w_i f_i(x)$, $w_i \geq 0, f_i(x)$ convex, then $f(x)$ is convex
- Convexity is preserved under a linear transformation
  - If $f(x) = g(Ax + b)$, g convex, then $f(x)$ is convex
- If f is a twice differentiable function of one variable, f is convex on an interval $[a, b] \subset \mathbb{R}$ iff (if and only if) its second derivative $f''(x) \geq 0$ in $[a, b]$

Convex Optimization

- How to solve

- ➢ No general way to solve
- ➢ Many algorithms developed for special classes of optimization problems (i.e., when $f(x)$ and $\mathcal{F}$ satisfy certain constraints):
  - ○ convex optimization problem (CO)
  - ○ linear program (LP)
  - ○ (mixed) integer linear program (MILP)
  - ○ quadratic program (QP), (Mixed) integer quadratic program (MIQP), semidefinite program (SDP), second-order cone program (SOCP), …
- ➢ Existing solvers and code packages for these problems

Convex Optimization

- How to solve

➢ Local search

➢ Iteratively improving an assignment

➢ Continuous and differentiable setting:

- Iteratively improving value of x
- Based on gradient

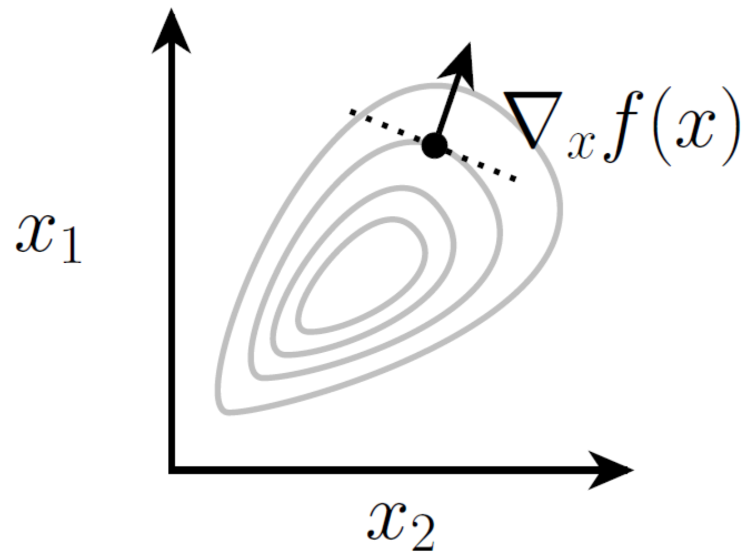
Convex Optimization

- How to solve

➢ For $f: \mathbb{R}^n \to \mathbb{R}$, **gradient** is the vector of partial derivative

- a multi-variable generalization of the derivative
- point in the direction of steepest increase in f

$$\nabla_x f(x) \in \mathbb{R}^n = \begin{bmatrix} \dfrac{\partial f(x)}{\partial x_1} \\ \dfrac{\partial f(x)}{\partial x_2} \\ \vdots \\ \dfrac{\partial f(x)}{\partial x_n} \end{bmatrix}$$



Convex Optimization

- How to solve

➢ Gradient descent: iteratively update the value of x

○ a simple algorithm for unconstrained optimization $\min_{x \in \mathbb{R}^n} f(x)$

Algorithm: Gradient Descent

Input: function f, initial point x_0, step size $\alpha > 0$

Initialize $x \leftarrow x_0$
Repeat
$\quad x \leftarrow x - \alpha \nabla_x f(x)$
Until convergence

○ Variants:

❑ How to choose x_0, e.g., $x_0 = 0$

❑ How to update α, e.g., $\alpha^{i+1} = \frac{(x^{i+1}-x^i)^T(\nabla_x f(x^{i+1}) - \nabla_x f(x^i))}{\|\nabla_x f(x^{i+1}) - \nabla_x f(x^i)\|_2^2}$

❑ How to define “convergence”, e.g., $\|x^{i+1} - x^i\|_2 \leq \epsilon$

Convex Optimization

• How to solve

➢ Projected Gradient Descent:
iteratively update the value of x while ensuring $x \in \mathcal{F}$

Algorithm: Projected Gradient Descent

Input: function f, initial point x_0, step size $\alpha > 0$

Initialize $x \leftarrow x_0$
Repeat
$\quad x \leftarrow P_{\mathcal{F}}(x - \alpha \nabla_x f(x))$
Until convergence

- $P_{\mathcal{F}}$ projects a point to the constraint set
- Variants

❑ How to choose $P_{\mathcal{F}}$, e.g., $P_{\mathcal{F}}(x) = \underset{x' \in \mathcal{F}}{\operatorname{argmin}} \|x - x'\|_2^2$

Convex Optimization

- How to solve

➢ Unconstrained and differentiable
- ○ gradient descent
- ○ set derivative to be 0:
- ❑ closed form solution
- ❑ Newton's method (if twice differentiable)

➢ Constrained and differentiable
- ○ projected gradient descent
- ○ interior point method

➢ Non-differentiable
- ○ ϵ-subgradient method
- ○ cutting plane method

Convex Optimization

- Apply

➢ Model a problem as a convex optimization problem
- define variable, feasible set, objective function
- prove it is convex (convex function + convex set)

➢ Solve the convex optimization problem
- Build up the model
- Call a solver, for example:
  - ❑ in R: CVXR
  - ❑ in Python: cvxpy, cvxopt
  - ❑ in MATLAB: fmincon, cvx

➢ Map the solution back to the original problem

Bibliography

- Hogben L. (Ed.) (2013).
  Handbook of Linear Algebra, 2nd edition,
  Chapman and Hall Book – CRC Press

- Boyd S. and L. Vandenberghe (2004).
  Convex Optimization,
  Cambridge University Press