Linear Algebra and Calculus

General notations

Vector: a vector is a list of numbers, usually written vertically as a column or horizontally as a row. The numbers that make up the list are called the entries of the vector.

Given a vector x with n entries, where each entry $\in x^i \in R$ represents the i^{th} component of the vector, it can be represented as:

$X = [x1,x2,x3,...,xn] \in R^n$

Matrix: a matrix is a rectangular array of numbers, usually written vertically as a column or horizontally as a row. The numbers that make up the list are called the entries of the matrix.

For a matrix A with m rows and n columns, we denote $A \in R^{m\times n}$. Each entry $A_{i,j} \in R$ in the matrix represents the element located in the i^{th} row and j^{th} column. The matrix A can be expressed as:

$A = [A1,1,A1,2,...,A1,n$

$A2,1,A2,2,...,A2,n$

$\dots\ Am,1,Am,2,...,Am,n] \in R^{m\times n}$

The vector x defined above can be considered as a n×1 matrix. It is specifically called a column vector and can be written as:

$X = [x1,x2,x3,...,xn] \in R^{n\times 1}$

Identity matrix

The identity matrix $I \in R^{n\times n}$ is a square matrix with ones on its main diagonal and zeroes everywhere else.

The matrix I is represented as:

$I = [1,0,0,...,0$

$0,1,0,...,0$

$\dots\ 0,0,0,...,1] \in R^{n\times n}$

Remark: For all matrices $A \in R^{n\times n}$, we have:

$A \times I = I \times A = A$

Matrix operations

Vector-vector multiplication- There are two types of vector-vector products:
For the inner product of vectors x and y:
Given $x,y \in R^n$ the inner product is given by:

$x.y = x_1y_1 + x_2y_2 + x_3y_3 + ... + x_ny_n$

This represents the summation of the products of corresponding elements of vectors x and y.

For the outer product of vectors x and y:
Given $x \in R^m$ and $y \in R^n$, the outer product is given by:

$x \otimes y = xy^T$

Where:

$$x \otimes y = [x_1y_1, x_1y_2, ..., x_1y_n \\ x_2y_1, x_2y_2, ..., x_2y_n \\ \ldots\ x_m, y_1, ..., x_mn]$$

This matrix is of size m×n and is formed by multiplying each element of x with each element of y.

Matrix-vector multiplication:

Given a matrix $A \in R^{m\times n}$ and a vector $x \in R^n$, the matrix-vector product is given by:
if have a matrix $A \in R^{m\times n}$ and a vector $x \in R^n$, their multiplication will result in a vector of size R^m.
The multiplication is defined as follows:

$y = A . x$
Here, y is the resulting vector $y \in R^m$ and its i^th element can be computed using:

$y_i = \mathrm{sum}(A_{i,j}.x_j)$ for j = 1 to n
Where y_i is the i^th element of vector y, $A_{i,j}$ is the element of the matrix A located in the i^th row and j^th column, and x_j
is the j^th element of vector x.

Matrix-matrix multiplication:

Given matrices $A \in R^{m\times n}$ and $B \in R^{n\times p}$, their product results in a matrix $C \in R^{m\times p}$:
$C = A . B$
The element $C_{ij} \in R$ is given by:
$C_{i,j} = \mathrm{sum}(A_{i,k}.B_{k,j})$ for k = 1 to n
Where $C_{i,j}$ is the element of matrix C located in the i^th row and j^th column, $A_{i,k}$ is the element of matrix A located in the i^th row and k^th column, and $B_{k,j}$ is the element of matrix B located in the k^th row and j^th column.

Transpose:

Given a matrix A, its transpose, denoted A^T, is obtained by flipping the matrix over its diagonal. This switches its row and column indices.

$(A^T)_{j,i} = A_{i,j}$ for all i and j

if A is of size $m \times n$, then A^T is of size $n \times m$.

Inverse:

Given an invertible square matrix A, its inverse is denoted A^{-1}. The unique property of the inverse matrix is:

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

Where I is the identity matrix of the same size as A.

Matrix calculus:

Gradient

Given a function $f:R^{m \times n} \to R$ and a matrix $A \in R^{m \times n}$, the gradient of f with respect to A is a matrix of size $m \times n$, denoted as $\nabla_A f(A)$. Each entry (i,j) of $\nabla_A f(A)$ corresponds to the partial derivative of f with respect to the (i,j)-th entry of A:

$[\nabla_A f(A)]_{i,j} = \partial f(A) / \partial A_{i,j}$

This means the entry at the i-th row and j-th column of the gradient matrix represents how f changes as $A_{i,j}$ changes, keeping all other entries of A constant.

Matrix properties

Norm

Given a vector space V, a norm is a function $N:V \to [0,+\infty)$ that satisfies the following properties for all vectors x,y in
V and scalar α:

1- Non-negativity: $N(x) \geq 0$ and $N(x)=0$ if and only if $x=0$.
2- Scalar multiplication: $N(\alpha x)=|\alpha|N(x)$.
3- Triangle inequality: $N(x+y) \leq N(x)+N(y)$.

These conditions ensure that the function N behaves like a measure of "length" or "size" for vectors in V.

Type of norms for vectors:

1 - L^1 norm (Manhattan):
$||x||_1 = |x1| + |x2| + ... + |xn|$

2- L^2 norm (Euclidean):
$||x||_2 = \sqrt{(|x1|^2 + |x2|^2 + ... + |xn|^2)}$

3- $L\infty$ norm (Maximum):
$||x||\infty = \max(|x1|,|x2|,...,|xn|)$

Type of norms for matrices:

1- Frobenius norm:
$||A||F = \sqrt{(\text{sum}(Ai,j^2))}$ for i = 1 to m and j = 1 to n