Review of Linear Algebra

Dr. Gerhard Roth

COMP 4102A

Winter 2014

Version 1

Linear algebra

- Is an important area of mathematics
- It is the basis of computer vision
- Is very widely taught, and there are many resources and books available
- We only focus on the basics
- Need to understand matrices, vectors and their basic operations

Linear Equations

A system of linear equations, e.g.

$$
\begin{aligned}
2x_1 + 4x_2 &= 2 \\
4x_1 + 11x_2 &= 1
\end{aligned}
$$

can be written in matrix form, m rows and n columns:

$$
\begin{bmatrix} 2 & 4 \\ 4 & 11 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 2 \\ 1 \end{bmatrix}
$$

or in general:

$$
Ax = b
$$

Matrix

A matrix is an $m \times n$ array of numbers.

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} = \left[a_{ij} \right]$$

Example:

$$A = \begin{bmatrix} 2 & 3 & 5 & 4 \\ -4 & 1 & 3 & 9 \\ 0 & 7 & 10 & 11 \end{bmatrix}$$

Matrix Arithmetic

Matrix addition

$$A_{m\times n} + B_{m\times n} = \left[a_{ij} + b_{ij}\right]_{m\times n}$$

Matrix multiplication

$$A_{m\times n} B_{n\times p} = C_{m\times p} \qquad c_{ij} = \sum_{k=1}^{n} a_{ik} b_{kj}$$

Matrix transpose

$$A^T = \left[a_{ji}\right]$$

$$(A+B)^T = A^T + B^T \qquad (AB)^T = B^T A^T$$

Multiplication not commutative

Matrix multiplication is not commutative

$$AB \neq BA$$

Example:

$$\begin{bmatrix} 2 & 5 \\ 3 & 1 \end{bmatrix}\begin{bmatrix} 6 & 2 \\ 1 & 5 \end{bmatrix} = \begin{bmatrix} 17 & 29 \\ 19 & 11 \end{bmatrix}$$

$$\begin{bmatrix} 6 & 2 \\ 1 & 5 \end{bmatrix}\begin{bmatrix} 2 & 5 \\ 3 & 1 \end{bmatrix} = \begin{bmatrix} 18 & 32 \\ 17 & 10 \end{bmatrix}$$

Symmetric Matrix

We say matrix A is symmetric if

$$A^T = A$$

Example:

$$A = \begin{bmatrix} 2 & 4 \\ 4 & 5 \end{bmatrix}$$

A symmetric matrix has to be a square matrix

Inverse of matrix

If A is a square matrix, the inverse of A, written A^{-1} satisfies:

$$AA^{-1} = I \qquad A^{-1}A = I$$

Where I, the identity matrix, is a diagonal matrix with all 1's on the diagonal.

$$I_2 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \qquad I_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Vectors – matrices with one column

$$x = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \qquad \text{e.g.} \qquad x = \begin{bmatrix} 2 \\ 3 \\ 5 \end{bmatrix}$$

A vector is an n by 1 matrix, it is a column in our book

The **length** or the **norm** of a vector is

$$\|x\| = \sqrt{x_1^2 + x_2^2 + \cdots + x_n^2}$$

e.g. $$\|x\| = \sqrt{2^2 + 3^2 + 5^2} = \sqrt{38}$$

Vector Arithmetic

Vector addition

$$u + v = \begin{bmatrix} u_1 \\ u_2 \end{bmatrix} + \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} = \begin{bmatrix} u_1 + v_1 \\ u_2 + v_2 \end{bmatrix}$$

Vector subtraction

$$u - v = \begin{bmatrix} u_1 \\ u_2 \end{bmatrix} - \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} = \begin{bmatrix} u_1 - v_1 \\ u_2 - v_2 \end{bmatrix}$$

Multiplication by scalar

$$\alpha u = \alpha \begin{bmatrix} u_1 \\ u_2 \end{bmatrix} = \begin{bmatrix} \alpha u_1 \\ \alpha u_2 \end{bmatrix}$$

Dot Product (inner product)

$$a = \begin{bmatrix} 2 \\ 3 \\ 5 \end{bmatrix} \qquad b = \begin{bmatrix} 4 \\ -3 \\ 2 \end{bmatrix}$$

$$a \cdot b = a^T b = \begin{bmatrix} 2 & 3 & 5 \end{bmatrix} \begin{bmatrix} 4 \\ -3 \\ 2 \end{bmatrix} = 2 \cdot 4 + 3 \cdot (-3) + 5 \cdot 2 = 9$$

$$a \cdot b = a^T b = a_1 b_1 + a_2 b_2 + \cdots + a_n b_n$$

Vectors: Dot Product (Inner product)

$$A \cdot B = A^T B \; = \begin{bmatrix} a & b & c \end{bmatrix} \begin{bmatrix} d \\ e \\ f \end{bmatrix} = ad + be + cf$$

Think of the dot product as a matrix multiplication

$$\|A\|^2 = A^T A = aa + bb + cc$$

The magnitude is the dot product of a vector with itself

$$A \cdot B = \|A\| \, \|B\| \cos(\theta)$$

The dot product is also related to the angle between the two vectors

Trace of Matrix

The trace of a matrix:

$$Tr(A) = \sum_{i=1}^{n} a_{ii}$$

Orthogonal Matrix

A matrix A is orthogonal if

$$A^T A = I \qquad \text{or} \qquad A^T = A^{-1}$$

Example:

$$A = \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix}$$

Matrix Transformation (and projections)

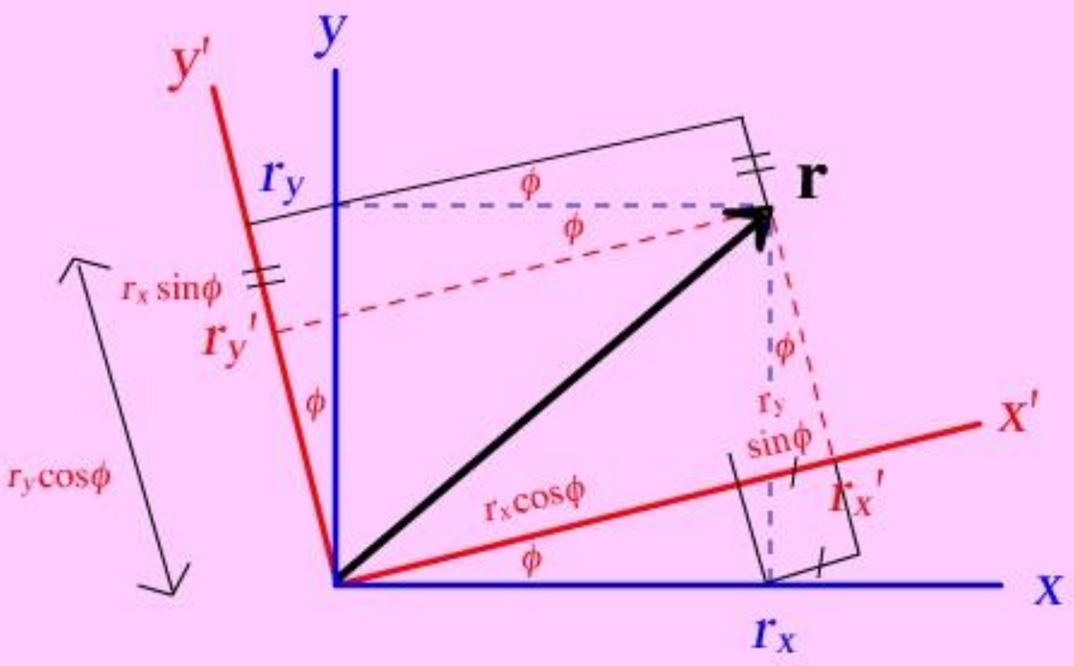
A matrix-vector multiplication transforms one vector to another

$$A_{m\times n}x_{n\times 1} = b_{m\times 1}$$

Example:

$$\begin{bmatrix} 2 & 5 \\ 3 & 1 \\ 4 & 2 \end{bmatrix}\begin{bmatrix} 3 \\ 7 \end{bmatrix} = \begin{bmatrix} 41 \\ 16 \\ 26 \end{bmatrix}$$

Coordinate Rotation

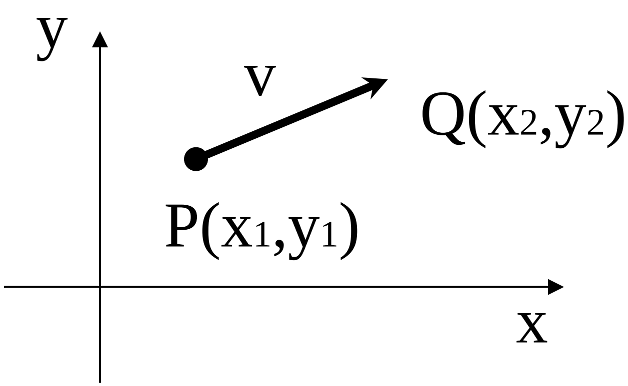


$$r_x' = r_x \cos\phi + r_y \sin\phi$$
$$r_y' = -r_x \sin\phi + r_y \cos\phi$$

$$\begin{bmatrix} r_x' \\ r_y' \end{bmatrix} = \begin{bmatrix} \cos\phi & \sin\phi \\ -\sin\phi & \cos\phi \end{bmatrix} \begin{bmatrix} r_x \\ r_y \end{bmatrix}$$

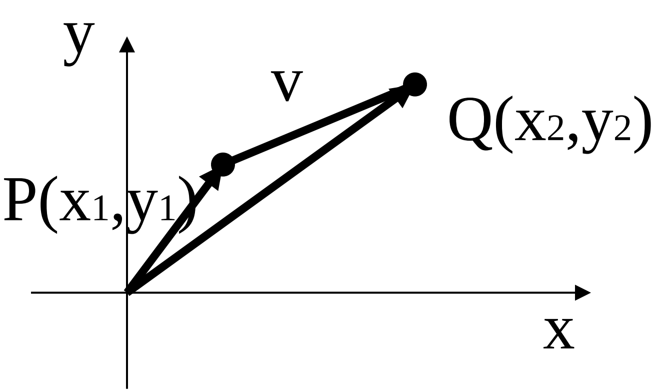
Vectors and Points

Two points in a Cartesian coordinate system define a vector



$$v = \begin{bmatrix} x_2 - x_1 \\ y_2 - y_1 \end{bmatrix}$$

A point can also be represented as a vector, defined by the point and the origin (0,0).



$$P = \begin{bmatrix} x_1 \\ y_1 \end{bmatrix} - \begin{bmatrix} 0 \\ 0 \end{bmatrix} = \begin{bmatrix} x_1 \\ y_1 \end{bmatrix} \qquad Q = \begin{bmatrix} x_2 \\ y_2 \end{bmatrix} - \begin{bmatrix} 0 \\ 0 \end{bmatrix} = \begin{bmatrix} x_2 \\ y_2 \end{bmatrix}$$

$$v = Q - P$$ or $$Q = P + v$$

Note: point and vector are different; vectors do not have positions

Least Squares

When m>n for an m-by-n matrix A, $Ax = b$ has no solution.

In this case, we look for an approximate solution.
We look for vector x such that

$$\|Ax - b\|^2$$

is as small as possible.

This is called the least squares solution.

Least Squares

Least squares solution of linear system of equations

$$Ax = b$$

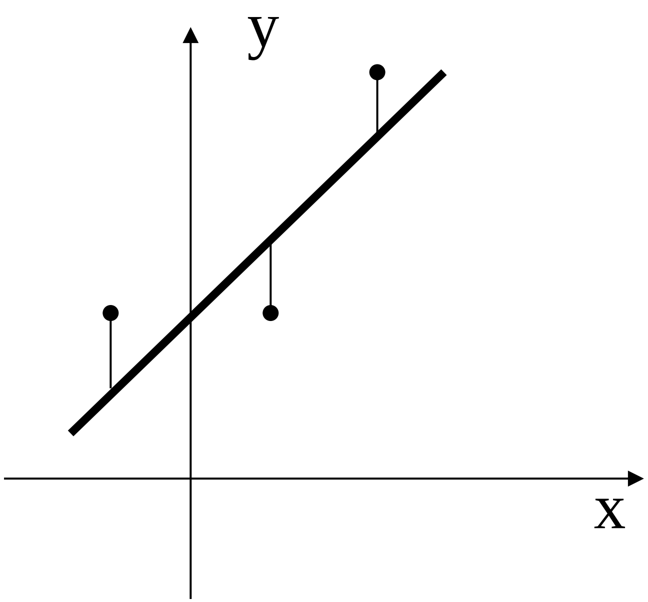
Normal equation: $A^T A x = A^T b$

$A^T A$ is square and symmetric

The Least Square solution $\bar{x} = (A^T A)^{-1} A^T b$

makes $\left\| A\bar{x} - b \right\|^2$ minimal.

Least Square Fitting of a Line



Line equations:

$$
\begin{aligned}
c + dx_1 &= y_1 \\
c + dx_2 &= y_2 \\
&\vdots \\
c + dx_m &= y_m
\end{aligned}
$$

$$
\begin{bmatrix} 1 & x_1 \\ 1 & x_2 \\ \vdots & \vdots \\ 1 & x_m \end{bmatrix} \begin{bmatrix} c \\ d \end{bmatrix} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_m \end{bmatrix}
$$

$$
Ax = y
$$

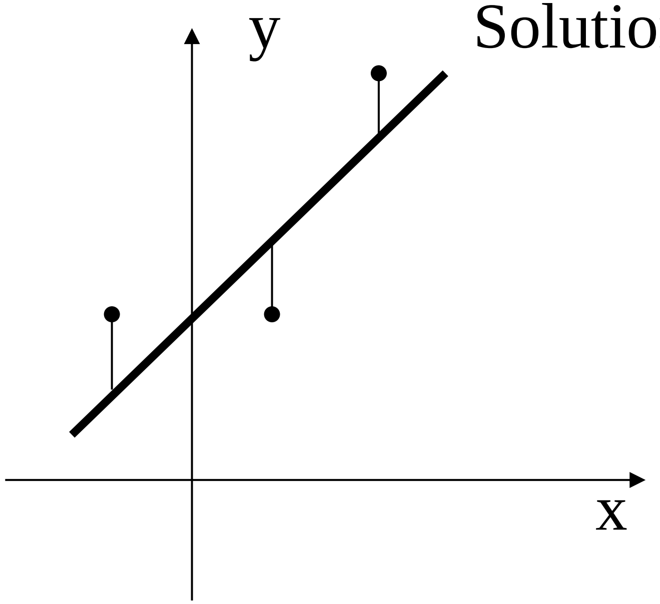
The best solution c, d is the one that minimizes:

$$
E^2 = \|y - Ax\|^2 = (y_1 - c - dx_1)^2 + \cdots + (y_m - c - dx_m)^2 .
$$

Least Square Fitting - Example

Problem: find the line that best fit these three points:

P1=(-1,1), P2=(1,1), P3=(2,3)



Solution:

$$c - d = 1$$
$$c + d = 1$$
$$c + 2d = 3$$

or

$$\begin{bmatrix} 1 & -1 \\ 1 & 1 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} c \\ d \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 3 \end{bmatrix}$$

$$A^T A x = A^T b$$

is

$$\begin{bmatrix} 3 & 2 \\ 2 & 6 \end{bmatrix} \begin{bmatrix} c \\ d \end{bmatrix} = \begin{bmatrix} 5 \\ 6 \end{bmatrix}$$

The solution is $c = \frac{9}{7}, d = \frac{4}{7}$ and best line is $\frac{9}{7} + \frac{4}{7} x = y$

Homogeneous System

- m linear equations with n unknowns A**x** = 0
- Assume that m >= n-1 and rank(A) = n-1
- Trivial solution is **x** = 0 but there are more
- If we have a given solution **x**, s.t. A**x** = 0 then c * **x** is also a solution since A(c* **x**) = 0
- Need to add a constraint on **x,**
  - Usually make **x** a unit vector $x^T x = 1$
- Can prove that the solution of A**x** = 0 satisfying this constraint is the eigenvector corresponding to the only zero eigenvalue of that matrix $A^T A$

Homogeneous System

- This solution can be computed using the eigenvector or SVD routine
  - Find the zero eigenvalue (or the eigenvalue almost zero)
  - Then the associated eigenvector is the solution **x**
- And any scalar times **x** is also a solution

Linear Independence

- A set of vectors is linear dependant if one of the vectors can be expressed as a linear combination of the other vectors.

$$v_k = \alpha_1 v_1 + \cdots + \alpha_{k-1} v_{k-1} + \alpha_{k+1} v_{k+1} + \cdots + \alpha_n v_n$$

- A set of vectors is linearly independent if none of the vectors can be expressed as a linear combination of the other vectors.

Eigenvalue and Eigenvector

We say that x is an eigenvector of a square matrix A if

$$Ax = \lambda x$$

λ is called eigenvalue and x is called eigenvector.

The transformation defined by A changes only the magnitude of the vector x

Example:

$$\begin{bmatrix} 3 & 2 \\ 1 & 4 \end{bmatrix}\begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 5 \\ 5 \end{bmatrix} = 5\begin{bmatrix} 1 \\ 1 \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} 3 & 2 \\ 1 & 4 \end{bmatrix}\begin{bmatrix} 2 \\ -1 \end{bmatrix} = \begin{bmatrix} 4 \\ -2 \end{bmatrix} = 2\begin{bmatrix} 2 \\ -1 \end{bmatrix}$$

5 and 2 are eigenvalues, and $\begin{bmatrix} 1 \\ 1 \end{bmatrix}$ and $\begin{bmatrix} 2 \\ -1 \end{bmatrix}$ are eigenvectors.

Properties of Eigen Vectors

- If $\lambda_1, \lambda_2, \ldots, \lambda_q$ are distinct eigenvalues of a matrix, then the corresponding eigenvectors $\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_q$ are linearly independent.

- A real, symmetric matrix has real eigenvalues with eigenvectors that can be chosen to be orthonormal.

SVD: Singular Value Decomposition

An $m \times n$ matrix A can be decomposed into:

$$A = UDV^T$$

U is $m \times m$, V is $n \times n$, both of them have orthogonal columns:

$$U^T U = I \qquad V^T V = I$$

D is an $m \times n$ diagonal matrix.

Example:

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

Singular Value Decomposition

- Any m by n matrix A can be written as product of three matrices $A = UDV^T$
- The columns of the m by m matrix U are mutually orthogonal unit vectors, as are the columns of the n by n matrix V
- The m by n matrix D is diagonal, and the diagonal elements, σ_i are called the singular values
- It is the case that $\sigma_1 \geq \sigma_2 \geq \ldots \sigma_n \geq 0$
- A matrix is non-singular if and only all of the singular values are not zero
- The condition number of the matrix is $\frac{\sigma_1}{\sigma_n}$
- If the condition number is large, then then matrix is almost singular and is called ill-conditioned

Singular Value Decomposition

- The rank of a square matrix is the number of linearly independent rows or columns
- For a square matrix (m = n) the number of non-zero singular values equals the rank of the matrix
- If A is a square, non-singular matrix, it's inverse can be written as $A^{-1} = VD^{-1}U^T$ where $A = UDV^T$
- The squares of the non zero singular values are the non-zero eigenvalues of both the n by n matrix $A^T A$ and of the m by m matrix AA^T
- The columns of U are the eigenvectors of AA^T
- The columns of V are the eigenvectors of $A^T A$