Maths for Signals and Systems

Linear Algebra in Engineering

Lectures 8-9, Friday 30th October 2015

DR TANIA STATHAKI

READER (ASSOCIATE PROFFESOR) IN SIGNAL PROCESSING
IMPERIAL COLLEGE LONDON

Semi-orthogonal matrices with more rows than columns

- The column vectors $q_1, \dots, q_n$ are **orthogonal** if ${q_i}^T \cdot q_j = 0$ for $i \neq j$.
- In order for a set of n vectors to satisfy the above, their dimension m must be at least n, i.e., $m \geq n$. This is because the maximum number of $m-$ dimensional vectors that can be orthogonal is m.
- If their lengths are all 1, then the vectors are called **orthonormal**.

$$ {q_i}^T \cdot q_j = \begin{cases} 0 & \text{when} \quad i \neq j \quad (\textbf{orthogonal} \text{ vectors}) \\ 1 & \text{when} \quad i = j \quad (\textbf{unit} \text{ vectors: } \|q_i\| = 1) \end{cases} $$

- I assign to a matrix with n orthonormal m −dimensional columns the special letter $Q_{m \times n}$.
- Now I will drop the subscript because no one uses it.
- I wish to deal first with the case where Q is strictly non-square (it is rectangular), and therefore, $m > n$.
- The matrix Q is called **semi-orthogonal**.

Semi-orthogonal matrices with more rows than columns

Problem:

Consider a semi-orthogonal matrix Q with real entries, where the number of rows m exceeds the number of columns n and the columns are orthonormal vectors.

Prove that $Q^T Q = I_{n\times n}$.

Solution:

$$Q^T Q = \begin{bmatrix} {q_1}^T \\ {q_2}^T \\ \vdots \\ {q_n}^T \end{bmatrix} \begin{bmatrix} q_1 & q_2 & \cdots & q_n \end{bmatrix} = I_{n\times n}.$$

- We see that Q^T is only an **inverse from the left**.
- This is because there isn't a matrix Q' for which $QQ' = I_{m\times m}$. This would imply that we could find m independent vectors of dimension n, with $m > n$. **This is not possible.**

Semi-orthogonal matrices: Generalization

- In linear algebra, a **semi-orthogonal matrix** is a non-square matrix with real entries where: if the number of rows exceeds the number of columns, then the columns are orthonormal vectors; but if the number of columns exceeds the number of rows, then the rows are orthonormal vectors.

- Equivalently, a rectangular matrix of dimension $m \times n$ is semi-orthogonal if

$$Q^T Q = I_{n\times n},\ m > n \text{ or } QQ^T = I_{m\times m},\ n > m$$

- The above formula yields the terms **left-invertible** or **right-invertible** matrix.

- In the above cases, the left or right inverse is the transpose of the matrix. For that reason, a rectangular orthogonal matrix is called **semi-unitary**. (To remind you: a **unitary** matrix is the one with an inverse being its transpose.)

Semi-orthogonal matrices: Generalization

Problem 1:

Show that for left-invertible, semi-orthogonal matrices of dimension $m \times n$, $m > n$ $\|Qx\| = \|x\|$ for every $n -$ dimensional vector x.

Solution:

$\|Qx\|^2 = (Qx)^T(Qx) = x^T Q^T Qx = x^T Ix = x^T x \Rightarrow \|Qx\|^2 = \|x\|^2 \Rightarrow \|Qx\| = \|x\|$.

Problem 2:

Show that for right-invertible, semi-orthogonal matrices of dimension $m \times n$, $m < n$, $\|Q^T x\| = \|x\|$ for every $m -$ dimensional vector x.

Solution:

$\|Q^T x\|^2 = (Q^T x)^T(Q^T x) = x^T QQ^T x = x^T Ix = x^T x \Rightarrow \|Q^T x\|^2 = \|x\|^2 \Rightarrow \|Q^T x\| = \|x\|$.

Orthogonal matrices

Problem 1:

Extend the relationship $Q^T Q = I_{n \times n}$ for the case when Q is a square matrix of dimension $n \times n$ and has orthogonal columns.

Solution:

$Q^T Q = I_{n \times n} \Rightarrow Q^{-1} = Q^T$. **The inverse is the transpose.**

Problem 2:

Prove that $QQ^T = I_{n \times n}$.

Solution:

Since Q is a full rank matrix we can find Q' such that $QQ' = I_{n \times n}$. This gives:

$$Q^T QQ' = Q^T I_{n \times n} \Rightarrow I_{n \times n}\, Q' = Q^T \Rightarrow Q' = Q^T$$

Therefore, we see that Q^T is the **two-sided inverse** of Q.

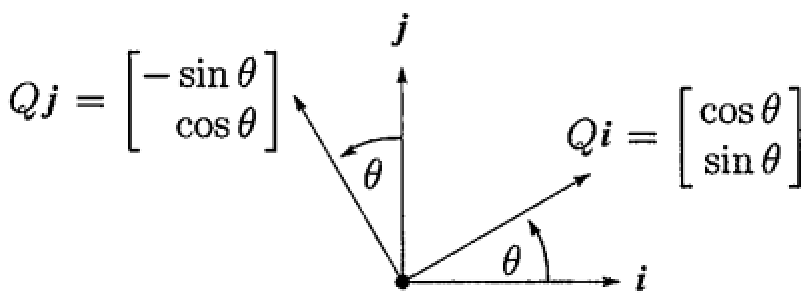
Examples of elementary orthogonal matrices Rotation

- **Rotation** matrix:

$Q = \begin{bmatrix} cos\theta & -sin\theta \\ sin\theta & cos\theta \end{bmatrix}$ and $Q^T = Q^{-1} = \begin{bmatrix} cos\theta & sin\theta \\ -sin\theta & cos\theta \end{bmatrix}$

Problem

- Show that the columns of Q are orthogonal (straightforward).
- Show that the columns of Q are unit vectors (straightforward).
- Explain the effect that the rotation matrix has on vectors $j = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$ and $i = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$, when it multiplies them from the left.
- The matrix causes rotation of the vectors.



Examples of elementary orthogonal matrices: Permutation

- **Permutation** matrices:

$Q = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix}$ and $Q = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$. $Q^T = Q^{-1}$ in both cases.

Problem

- Show that the columns of Q are orthogonal (straightforward).
- Show that the columns of Q are unit vectors (straightforward).
- Explain the effect that the permutation matrices have on a random vector $\begin{bmatrix} x \\ y \\ z \end{bmatrix}$ or $\begin{bmatrix} x \\ y \end{bmatrix}$ when they multiply the vector from the left.
- The matrices cause re-ordering of the elements of these vectors.

Examples of elementary orthogonal matrices Householder Reflection

- **Householder Reflection** matrices

  $Q = I - 2uu^T$ with u any vector that satisfies the condition $\|u\|_2 = 1$ (unit vector).
  $Q^T = I^T - (2uu^T)^T = I - 2uu^T = Q$
  $Q^T Q = Q^2 = I$

  Problem

  - For $u_1 = [1 \quad 0]^T$ and $u_2 = \left[1/\sqrt{2} \quad -1/\sqrt{2}\right]^T$ find $Q_i = I - 2u_i {u_i}^T$, $i = 1,2$.
  - Explain the effect that matrix Q_1 has on the vector $\begin{bmatrix} x \\ y \end{bmatrix}$ when it multiplies the vector from the left.
  - Explain the effect that matrix Q_2 has on the vector $\begin{bmatrix} x \\ y \end{bmatrix}$ when it multiplies the vector from the left.

- A generalized definition is $Q = I - 2\frac{vv^T}{\|v\|^2}$ with v any column vector.

Imperial College London

The Gram-Schmidt process

- The goal here is to start with three independent vector a, b, c and construct three orthogonal vectors A, B, C and finally three orthonormal vectors.

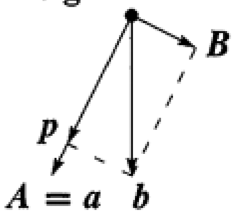
$$q_1 = A/\|A\|,\ q_2 = B/\|B\|,\ q_3 = C/\|C\|$$

- We begin by choosing $A = a$. This first direction is accepted.
- The next direction B must be perpendicular to A. Start with b and subtract its projection along A. This leaves the perpendicular part, which is the orthogonal vector B (what we knew before as error!), defined as:

$$B = b - \frac{AA^T}{A^TA}b$$

Problem: Show that A and B are orthogonal.

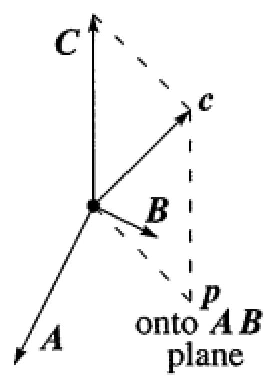
Problem: Show that if a and b are independent then B is not zero.



The Gram-Schmidt process

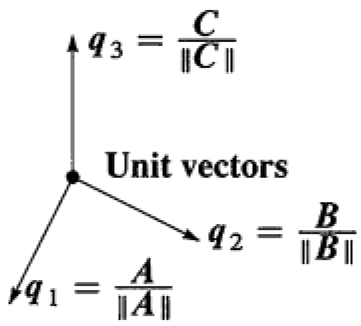
- The third direction starts with c. This is not a combination of A and B.
- Most likely c is not perpendicular to A and B .
- Therefore, subtract its components in those two directions to get C:

$$C = c - \frac{AA^T}{A^TA}c - \frac{BB^T}{B^TB}c$$



The Gram-Schmidt process: Generalization

- **In general we subtract from every new vector its projections in the directions already set.**
- If we had a fourth vector d, we would subtract three projections onto A, B, C to get D.
- We make the resulting vectors orthonormal.
- This is done by dividing the vectors with their magnitudes.



The factorization $A = QR$ (QR decomposition)

- Assume matrix A whose columns are a, b, c.
- Assume matrix Q whose columns are q_1, q_2, q_3 defined previously.
- We are looking for a matrix R such that $A = QR$. Since Q is an orthogonal matrix we have that $R = Q^T A$.

$$R = Q^T A = \begin{bmatrix} {q_1}^T \\ {q_2}^T \\ {q_3}^T \end{bmatrix} \begin{bmatrix} a & b & c \end{bmatrix} = \begin{bmatrix} {q_1}^T a & {q_1}^T b & {q_1}^T c \\ {q_2}^T a & {q_2}^T b & {q_2}^T c \\ {q_3}^T a & {q_3}^T b & {q_3}^T c \end{bmatrix}$$

- We know that from the method that was used to construct q_i we have

$$ {q_2}^T a = 0,\ \ {q_3}^T a = 0,\ \ {q_3}^T b = 0$$

and therefore,

$$R = \begin{bmatrix} {q_1}^T a & {q_1}^T b & {q_1}^T c \\ 0 & {q_2}^T b & {q_2}^T c \\ 0 & 0 & {q_3}^T c \end{bmatrix}$$

- QR decomposition can facilitate the solution of the system $Ax = b$, since $Ax = b \Rightarrow QRx = b \Rightarrow Rx = Q^T b$. The later system is easy to solve due to the upper triangular form of R.
- **So far you have learnt two types of decompositions: the LU and the QR.**

Determinants

- The **Determinant** is a crucial number associated with square matrices only.
- It is denoted by $\det(A) = |A|$. These are two different symbols we use for determinants.
- If a matrix A is invertible, that means $\det(A) \neq 0$.
- Furthermore, $\det(A) \neq 0$ means that matrix A is invertible.
- For a 2×2 matrix $\begin{bmatrix} a & b \\ c & d \end{bmatrix}$ the determinant is defined as $\begin{vmatrix} a & b \\ c & d \end{vmatrix} = ad - bc$. This formula is explicitly associated with the solution of the system $Ax = b$ where A is a 2×2 matrix.

Properties of determinants

1. $\det(I) = 1$. This is easy to show in the case of a 2×2 matrix using the formula of the previous slide.

2. If we exchange two rows of a matrix the sign of the determinant reverses. Therefore:
   - If we perform an even number of row exchanges the determinant remains the same.
   - If we perform an odd number of row exchanges the determinant changes sign.
   - Hence, the determinant of a **permutation** matrix is 1 or -1.

$\begin{vmatrix} 1 & 0 \\ 0 & 1 \end{vmatrix} = 1$ and $\begin{vmatrix} 0 & 1 \\ 1 & 0 \end{vmatrix} = -1$ as expected.

Properties of determinants

3a. If a row is multiplied with a scalar, the determinant is multiplied with that scalar too, i.e., $\begin{vmatrix} ta & tb \\ c & d \end{vmatrix} = t\begin{vmatrix} a & b \\ c & d \end{vmatrix}$.

3b. $\begin{vmatrix} a+a' & b+b' \\ c & d \end{vmatrix} = \begin{vmatrix} a & b \\ c & d \end{vmatrix} + \begin{vmatrix} a' & b' \\ c & d \end{vmatrix}$

Note that $\det(A+B) \neq \det(A) + \det(B)$

I observe linearity only for a single row.

4. Two equal rows leads to $det = 0$.

- As mentioned, if I exchange rows the sign of the determinant changes.
- In that case the matrix is the same and therefore, the determinant should remain the same.
- Therefore, the determinant must be zero.
- This is also expected from the fact that the matrix is not invertible.

Properties of determinants

5. $\begin{vmatrix} a & b \\ c-la & d-lb \end{vmatrix} = \begin{vmatrix} a & b \\ c & d \end{vmatrix} + \begin{vmatrix} a & b \\ -la & -lb \end{vmatrix} = \begin{vmatrix} a & b \\ c & d \end{vmatrix} - l \begin{vmatrix} a & b \\ a & b \end{vmatrix} = \begin{vmatrix} a & b \\ c & d \end{vmatrix}$
   Therefore, the determinant after elimination remains the same.

6. A row of zeros leads to $det = 0$. This can verified as follows for any matrix:
$$\begin{vmatrix} 0 & 0 \\ c & d \end{vmatrix} = \begin{vmatrix} 0 \cdot a & 0 \cdot b \\ c & d \end{vmatrix} = 0 \begin{vmatrix} a & b \\ c & d \end{vmatrix} = 0$$

7. Consider an upper triangular matrix ($*$ is a random element)
$$\begin{vmatrix} d_1 & * & \dots & * \\ 0 & d_2 & \dots & * \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & d_n \end{vmatrix} = d_1 d_2 \dots d_n$$
   I can easily show the above using the following steps:
   - I transform the upper triangular matrix to a diagonal one using elimination.
   - I use property 3a n times.
   - I end up with the determinant $\prod_{i=1}^{n} d_i \det(I) = \prod_{i=1}^{n} d_i$.
   - Same comments are valid for a lower triangular matrix.

Properties of determinants

8. $\det(A) = 0$ when A is singular. This is because if A is singular I get a row of zeros by elimination.

   Using the same concept I can say that if A is invertible then $\det(A) \neq 0$.

   In general I have $A \to U \to D$, $\det(A) = d_1 d_2 \dots d_n$ =product of pivots.

9. $\det(AB) = \det(A)\det(B)$

   $\det(A^{-1}) = \frac{1}{\det(A)}$

   $\det(A^2) = [\det(A)]^2$

   $\det(2A) = 2^n \det(A)$ where $A: n \times n$

10. $\det(A^T) = \det(A)$.
    - In order to show that, we use the LU decomposition of A and the above properties. $A = LU$ and therefore $A^T = U^T L^T$. Determinant is always product of pivots.
    - This property can also be proved by the use of **induction**.

Determinant of a 2×2 matrix

- The goal is to find the determinant of a 2×2 matrix $\begin{vmatrix} a & b \\ c & d \end{vmatrix}$ using the properties described previously.
- We know that $\begin{vmatrix} 1 & 0 \\ 0 & 1 \end{vmatrix} = 1$ and $\begin{vmatrix} 0 & 1 \\ 1 & 0 \end{vmatrix} = -1$.
- $\begin{vmatrix} a & b \\ c & d \end{vmatrix} = \begin{vmatrix} a & 0 \\ c & d \end{vmatrix} + \begin{vmatrix} 0 & b \\ c & d \end{vmatrix} = \begin{vmatrix} a & 0 \\ c & 0 \end{vmatrix} + \begin{vmatrix} a & 0 \\ 0 & d \end{vmatrix} + \begin{vmatrix} 0 & b \\ c & 0 \end{vmatrix} + \begin{vmatrix} 0 & b \\ 0 & d \end{vmatrix} =$
  $0 + \begin{vmatrix} a & 0 \\ 0 & d \end{vmatrix} + \begin{vmatrix} 0 & b \\ c & 0 \end{vmatrix} + 0 = ad \begin{vmatrix} 1 & 0 \\ 0 & 1 \end{vmatrix} + bc \begin{vmatrix} 0 & 1 \\ 1 & 0 \end{vmatrix} = ad - bc$
- I can realize the above analysis for 3×3 matrices.
- I break the determinant of a 2×2 random matrix into 4 determinants of simpler matrices.
- In the case of a 3×3 matrix I break it into 27 determinants.
- And so on.

Determinant of any matrix

- For the case of a 2×2 matrix $\begin{vmatrix} a & b \\ c & d \end{vmatrix}$ we got:

$$\begin{vmatrix} a & b \\ c & d \end{vmatrix} = \begin{vmatrix} a & 0 \\ c & 0 \end{vmatrix} + \begin{vmatrix} a & 0 \\ 0 & d \end{vmatrix} + \begin{vmatrix} 0 & b \\ c & 0 \end{vmatrix} + \begin{vmatrix} 0 & b \\ 0 & d \end{vmatrix} = 0 + \begin{vmatrix} a & 0 \\ 0 & d \end{vmatrix} + \begin{vmatrix} 0 & b \\ c & 0 \end{vmatrix} + 0$$

- The determinants which survive have strictly one entry from each row and each column.
- The above is a universal conclusion.

Determinant of any matrix

- For the case of a 3×3 matrix $\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}$ we got:

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = \begin{vmatrix} a_{11} & 0 & 0 \\ 0 & a_{22} & 0 \\ 0 & 0 & a_{33} \end{vmatrix} + \begin{vmatrix} a_{11} & 0 & 0 \\ 0 & 0 & a_{23} \\ 0 & a_{32} & 0 \end{vmatrix} + \cdots =$$

$$a_{11}a_{22}a_{33} - a_{11}a_{23}\, a_{32} + \cdots$$

- As mentioned the determinants which survive have strictly one entry from each row and each column.

Determinant of any matrix

- For the case of a 2×2 matrix the determinant has 2 survived terms.
- For the case of a 3×3 matrix the determinant has 6 survived terms.
- For the case of a 4×4 matrix the determinant has 24 survived terms.
- For the case of a $n \times n$ matrix the determinant has $n!$ survived terms.
  - The elements from the first row can be chosen in n different ways.
  - The elements from the second row can be chosen in $(n-1)$ different ways.
  - and so on…

Problem

Find the determinant of the following matrix:

$$\begin{bmatrix} 0 & 0 & 1 & 1 \\ 0 & 1 & 1 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 0 & 0 & 1 \end{bmatrix}$$

Big Formula for the determinant

- For the case of a $n \times n$ matrix the determinant has $n!$ terms.

$$\det(A) = \sum_{n!\text{terms}} \pm a_{1a} a_{2b} a_{3c} \dots a_{nz}$$

  - ➢ $a, b, c, \dots, z$ are different columns.
  - ➢ In the above summation, half of the terms have a plus and half of them have a minus sign.

Big Formula for the determinant

- For the case of a $n \times n$ matrix, **cofactors** consist of a method which helps us to connect a determinant to determinants of smaller matrices.

$$\det(A) = \sum_{n!\text{terms}} \pm a_{1a} a_{2b} a_{3c} \dots a_{nz}$$

- For a 3×3 matrix we have $\det(A) = a_{11}(a_{22}a_{33} - a_{23}\, a_{32}) + \cdots$
- $a_{22}a_{33} - a_{23}\, a_{32}$ is the determinant of a 2×2 matrix which is a sub-matrix of the original matrix.

Cofactors

- The **cofactor** of element a_{ij} is defined as follows:

$$C_{ij} = \pm\det\left[(n-1)\times(n-1)\text{ matrix } A_{ij}\right]$$

  - A_{ij} is the $(n-1)\times(n-1)$ that is obtained from the original matrix if row i and column j are eliminated.
  - We keep the $+$ if $(i+j)$ is even.
  - We keep the $-$ if $(i+j)$ is odd.

- Cofactor formula along row 1:

$$\det(A) = a_{11}C_{11} + a_{12}C_{12} + \cdots + a_{1n}\,C_{1n}$$

- Generalization:
  - Cofactor formula along row i: $\det(A) = a_{i1}C_{i1} + a_{i2}C_{i2} + \cdots + a_{in}\,C_{in}$
  - Cofactor formula along column j: $\det(A) = a_{1j}C_{1j} + a_{2j}C_{2j} + \cdots + a_{nj}\,C_{nj}$

- Cofactor formula along any row or column can be used for the final estimation of the determinant.

Estimation of the inverse A^{-1} using cofactors

- For a 2×2 matrix it is quite easy to show that

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

- Big formula for A^{-1}

$$A^{-1} = \frac{1}{\det(A)} C^T$$

$$AC^T = \det(A) \cdot I$$

- C_{ij} is the cofactor of a_{ij} which is a sum of products of $(n-1)$ entries.
- In general

$$\begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn} \end{bmatrix} \begin{bmatrix} C_{11} & \cdots & C_{n1} \\ \vdots & & \vdots \\ C_{1n} & \cdots & C_{nn} \end{bmatrix} = \det(A) \cdot I$$

Solve $Ax = b$ when A is square and invertible

- The solution of the system $Ax = b$ when A is square and invertible can be now obtained from

$$x = A^{-1}b = \frac{1}{\det(A)} C^T b$$

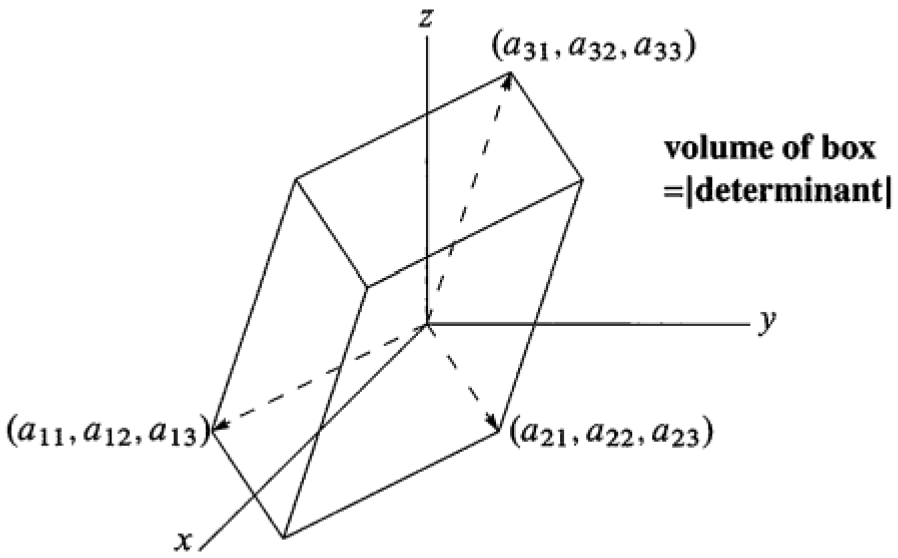
- **Cramer's rule:**
  - First element of vector x is $x_1 = \frac{\det(B_1)}{\det(A)}$. Then $x_2 = \frac{\det(B_2)}{\det(A)}$ and so on.
  - What are these matrices B_i?

$$B_1 = [b \vdots \quad \text{last } (n-1) \text{ columns of } A]$$

  - B_1 is obtained by A if we replace the first column with b. B_i is obtained by A if we replace the ith column with b.
  - In practice we must find $(n+1)$ determinants.

The determinant is the volume of a box

- Consider A to be a matrix of size 3×3.
- Observe the three-dimensional box (parallelepiped) formed from the three rows or columns of A.
- It can be proven that $\mathrm{abs}(|A|)$ ≡volume of the box.



$\det(A) \equiv$ volume of a box

- Take $A = I$. Then the box mentioned previously is the unit cube and its volume is 1.

  Problem:

  Consider an orthogonal square matrix Q. Prove that $\det(Q) = 1$ or -1.

  Solution:

  $Q^T Q = I \Rightarrow \det(Q^T Q) = \det(Q^T)\det(Q) = \det(I) = 1$

  But $\det(Q^T) = \det(Q) \Rightarrow |Q|^2 = 1 \Rightarrow |Q| = \pm 1$

- The above results verifies also the fact that the determinant of a 3×3 matrix is the volume of the cube that is formed by the rows of the matrix. This is because if you consider $A = Q$ with Q being an orthogonal matrix, the related box is a rotated version of the unit cube in the 3D space. Its volume is again 1.

- Take Q and double one of its vectors. The cube's volume doubles (you have two cubes sitting on top of each other.) The determinant doubles as well (property 3a).