FUNDAMENTALS OF LEARNING AND INFORMATION PROCESSING

SESSION 01: LINEAR ALGEBRA (I)

Dr Gang Li

Deakin University, Geelong, Australia

2018-11-06

Table of Content

Recommended Resources

! *Linear Algebra* is essential for data mining, machine learning and pattern recognition. I will try to emphasize the most popular linear algebra techniques as used in those areas. In order to know them inside-out, upside-down, you should:

- get an excellent textbook, at the right level
- do all exercises in the textbook

G. Strang. *Linear Algebra and Its Applications*. Brooks/Cole, 2006
G. Strang. MIT linear algebra 18.06, 2005
T. Tao. UCLA Linear Algebra Math115A, 2002

Linear Model $Ax = b$

Problem

Consider problems of the following form: given a number of measurements $A = [a_1, \cdots, a_n]$ and an outcome B, build a linear model which uses the measurements A to predict the outcome B.

- A = clinical measurements of patients, B = level of cancer specific antigen.
- A = atmospheric measurements , B = occurrence of PM25.

a	1	2	3	4	5
b	7.97	10.2	14.2	16.0	21.2

- We wish to find x_0 and x_1 such that $x_0 + x_1 \alpha = \beta$. Thus

$$\begin{aligned} x_0 + x_1 &= 7.97 \\ x_0 + 2x_1 &= 10.2 \\ x_0 + 3x_1 &= 14.2 \\ x_0 + 4x_1 &= 16.0 \\ x_0 + 5x_1 &= 21.2 \end{aligned}$$

- We wish to find x_0 and x_1 such that $x_0 + x_1 \alpha = \beta$. Thus in matrix form:

$$\begin{bmatrix} 1 & 1 \\ 1 & 2 \\ 1 & 3 \\ 1 & 4 \\ 1 & 5 \end{bmatrix} \begin{bmatrix} x_0 \\ x_1 \end{bmatrix} = \begin{bmatrix} 7.97 \\ 10.2 \\ 14.2 \\ 16.0 \\ 21.2 \end{bmatrix}$$

- In general, the model $\beta = x_0 + \sum_{j=1}^{n} x_j \alpha_j$ can be written in the form

$$Ax = b, \quad x = (x_0 x_1 \cdots x_n)^T$$

To do this, one must have $A = [a_0 a_1 \cdots a_n]$, where a_0 is a vector of ones, a_i are the vectors for the corresponding indicators; b is the vector for the outputs.

- Fitting a linear model to data is essential in *linear algebra*, as well as *probability*.

Model

Let $A \in \mathbb{R}^{m \times n}$, $x \in \mathbb{R}^n$ and $b \in \mathbb{R}^m$. The system is

$$Ax = b$$

Matrix Multiplication

- Matrix is a rectangular array of data, with elements are real numbers (in this FLIP course).
- Vector is in general a column vector (in this FLIP course).

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} \in \mathbb{R}^{m \times n}$$

- Matrix-Vector Multiplication

$$Ax = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \cdots \\ x_n \end{bmatrix} = \begin{bmatrix} \sum_{j=1}^{n} a_{1j} x_j \\ \sum_{j=1}^{n} a_{2j} x_j \\ \cdots \\ \sum_{j=1}^{n} a_{mj} x_j \end{bmatrix}$$

$$\begin{bmatrix} \leftarrow & \cdots & \rightarrow \\ \leftarrow & \cdots & \rightarrow \\ \leftarrow & \cdots & \rightarrow \end{bmatrix} \begin{bmatrix} \uparrow \\ \vdots \\ \downarrow \end{bmatrix} = \begin{bmatrix} \times \\ \vdots \\ \times \end{bmatrix}$$

- Vector-Matrix Multiplication

$$xA = \begin{bmatrix} x_1 & x_2 & \cdots & x_m \end{bmatrix} \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} = \begin{bmatrix} \sum_{i=1}^{m} x_i a_{i1} & \sum_{i=1}^{m} x_i a_{i2} & \cdots & \sum_{i=1}^{m} x_i a_{in} \end{bmatrix}$$

$$\begin{bmatrix} \leftarrow & \cdots & \rightarrow \end{bmatrix} \begin{bmatrix} \uparrow & \uparrow & \vdots & \uparrow \\ | & | & \vdots & | \\ \downarrow & \downarrow & \vdots & \downarrow \end{bmatrix} = \begin{bmatrix} \times & \times & \cdots & \times \end{bmatrix}$$

Matrix Multiplication

Info

Alternative presentation of *Matrix-Vector* multiplication:
a *linear combination* (*weighted combination*) of the columns of A.

- Denote the *column vectors* of the matrix A by a_j. Then

$$y = Ax = \begin{bmatrix} a_1 & a_2 & \cdots & a_n \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \sum_{j=1}^{n} x_j a_j$$

- For example

$$\begin{bmatrix} 2 & 3 \\ 6 & 4 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 5 \\ -2 \end{bmatrix} = 5 \cdot \begin{bmatrix} 2 \\ 6 \\ 1 \end{bmatrix} - 2 \begin{bmatrix} 3 \\ 4 \\ 0 \end{bmatrix} = \begin{bmatrix} 4 \\ 22 \\ 5 \end{bmatrix}$$

Matrix Multiplication

Info

Alternative presentation of *Vector-Matrix* multiplication:
a *linear combination* (*weighted combination*) of the rows of A.

- Denote the *row vectors* of the matrix A by a_i. Then

$$y = xA = \begin{bmatrix} x_1 & x_2 & \cdots & x_m \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_m \end{bmatrix} = \sum_{i=1}^{m} x_i a_i$$

- For example

$$\begin{bmatrix} 5 & -2 & 0 \end{bmatrix} \begin{bmatrix} 2 & 3 \\ 6 & 4 \\ 1 & 0 \end{bmatrix} = 5 \cdot \begin{bmatrix} 2 & 3 \end{bmatrix} - 2 \begin{bmatrix} 6 & 4 \end{bmatrix} + 0 \begin{bmatrix} 1 & 0 \end{bmatrix} = \begin{bmatrix} -2 & 7 \end{bmatrix}$$

Matrix Multiplication

Info Alternative presentation of *Matrix-Matrix* multiplication:

- For two matrix $A \in \mathbb{R}^{m\times n}$, and $B \in \mathbb{R}^{n\times p}$, the multiplication of them can be considered as the sum of *columns* of A × *rows* of B.
  - $v_i v_j^T$ is a full size matrix, called the *outer product* of vector v_i and v_j.
  - $v_i^T v_j$ is a *scalar*, called the *inner product* of vector v_i and v_j.

$$AB = \begin{bmatrix} a_1 & a_2 & \cdots & a_n \end{bmatrix} \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix} = \sum_{i=1}^{m} a_i b_i$$

- For example

$$\begin{bmatrix} 2 & 7 \\ 3 & 8 \\ 4 & 9 \end{bmatrix} \begin{bmatrix} 1 & 6 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} 2 \\ 3 \\ 4 \end{bmatrix} \cdot \begin{bmatrix} 1 & 6 \end{bmatrix} + \begin{bmatrix} 7 \\ 8 \\ 9 \end{bmatrix} \cdot \begin{bmatrix} 0 & 0 \end{bmatrix} = \begin{bmatrix} 2 & 12 \\ 3 & 18 \\ 4 & 24 \end{bmatrix}$$

Matrix Multiplication

Info Alternative presentation of *Matrix-Matrix* multiplication:

- For two matrix $A \in \mathbb{R}^{m\times n}$, and $B \in \mathbb{R}^{n\times p}$, the multiplication of them can be considered as the row-wise combination of (*rows* of matrix A × matrix B).

$$AB = \begin{bmatrix} a_1 \\ a_2 \\ \cdots \\ a_m \end{bmatrix} B = \begin{bmatrix} a_1 B \\ a_2 B \\ \cdots \\ a_m B \end{bmatrix}$$

- For example

$$\begin{bmatrix} 2 & 7 \\ 3 & 8 \\ 4 & 9 \end{bmatrix} \begin{bmatrix} 1 & 6 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} \begin{bmatrix} 2 & 7 \end{bmatrix} \begin{bmatrix} 1 & 6 \\ 0 & 0 \end{bmatrix} \\ \begin{bmatrix} 3 & 8 \end{bmatrix} \begin{bmatrix} 1 & 6 \\ 0 & 0 \end{bmatrix} \\ \begin{bmatrix} 4 & 9 \end{bmatrix} \begin{bmatrix} 1 & 6 \\ 0 & 0 \end{bmatrix} \end{bmatrix} = \begin{bmatrix} 2 & 12 \\ 3 & 18 \\ 4 & 24 \end{bmatrix}$$

Matrix Multiplication

Info Alternative presentation of *Matrix-Matrix* multiplication:

- For two matrix $A \in \mathbb{R}^{m\times n}$, and $B \in \mathbb{R}^{n\times p}$, the multiplication of them can be considered as the column-wise combination of (matrix A × columns of matrix B)

$$AB = A\begin{bmatrix} b_1 & b_2 & \cdots & b_p \end{bmatrix} = \begin{bmatrix} Ab_1 & Ab_2 & \cdots & Ab_p \end{bmatrix}$$

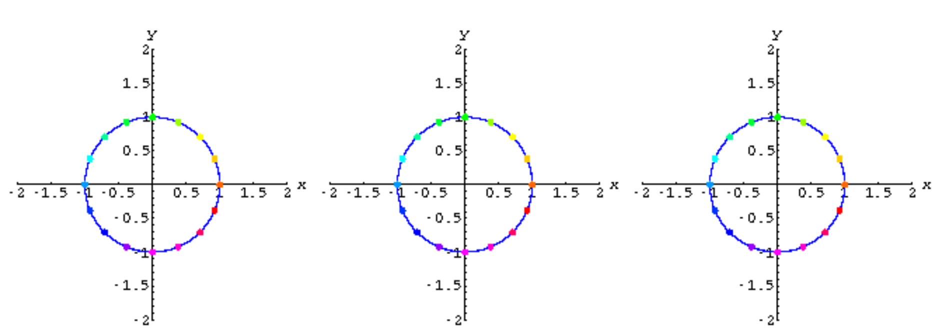
- For example

$$\begin{bmatrix} 2 & 7 \\ 3 & 8 \\ 4 & 9 \end{bmatrix}\begin{bmatrix} 1 & 6 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} \left[\begin{smallmatrix} 2 & 7 \\ 3 & 8 \\ 4 & 9 \end{smallmatrix}\right]\left[\begin{smallmatrix} 1 \\ 0 \end{smallmatrix}\right] & \left[\begin{smallmatrix} 2 & 7 \\ 3 & 8 \\ 4 & 9 \end{smallmatrix}\right]\left[\begin{smallmatrix} 6 \\ 0 \end{smallmatrix}\right] \end{bmatrix} = \begin{bmatrix} 2 & 12 \\ 3 & 18 \\ 4 & 24 \end{bmatrix}$$

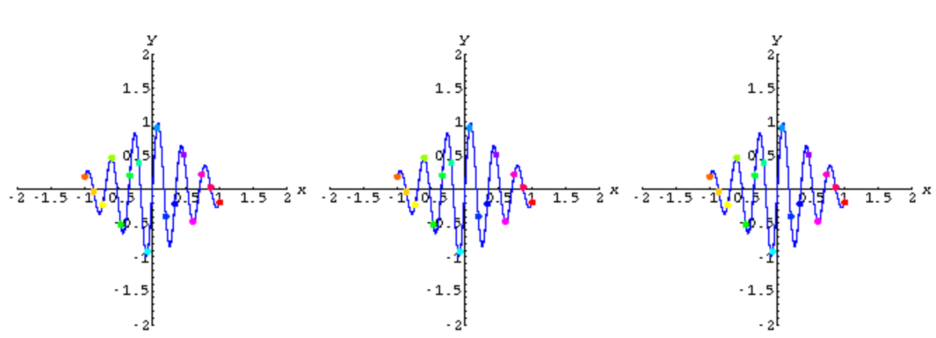
Matrix Hits

Defn When you *hit* a point $x \in \mathbb{R}^m$ with a matrix $A \in \mathbb{R}^{m\times m}$, you get another point Ax.

- For example, a square matrix $\left[\begin{smallmatrix} 1 & 3 \\ -3 & 2 \end{smallmatrix}\right]$ hits $\left[\begin{smallmatrix} 3 \\ 1 \end{smallmatrix}\right]$, you get $\left[\begin{smallmatrix} 1 & 3 \\ -3 & 2 \end{smallmatrix}\right]\left[\begin{smallmatrix} 3 \\ 1 \end{smallmatrix}\right] = \left[\begin{smallmatrix} 6 \\ -7 \end{smallmatrix}\right]$
- What if we *hit* a set of points with a matrix?
  - For example, use matrix $\left[\begin{smallmatrix} 0.26 & 0.68 \\ 0.68 & 1.7 \end{smallmatrix}\right]$, $\left[\begin{smallmatrix} 1.6 & -0.65 \\ -0.65 & -0.3 \end{smallmatrix}\right]$, and $\left[\begin{smallmatrix} 1.4 & -0.62 \\ -1.1 & -1.7 \end{smallmatrix}\right]$ to *hit* a *unit circle* respectively.



  - For example, use matrix $\left[\begin{smallmatrix} 0.26 & 0.68 \\ 0.68 & 1.7 \end{smallmatrix}\right]$, $\left[\begin{smallmatrix} 1.6 & -0.65 \\ -0.65 & -0.3 \end{smallmatrix}\right]$, and $\left[\begin{smallmatrix} 1.4 & -0.62 \\ -1.1 & -1.7 \end{smallmatrix}\right]$ to *hit* a *sine wave* respectively.

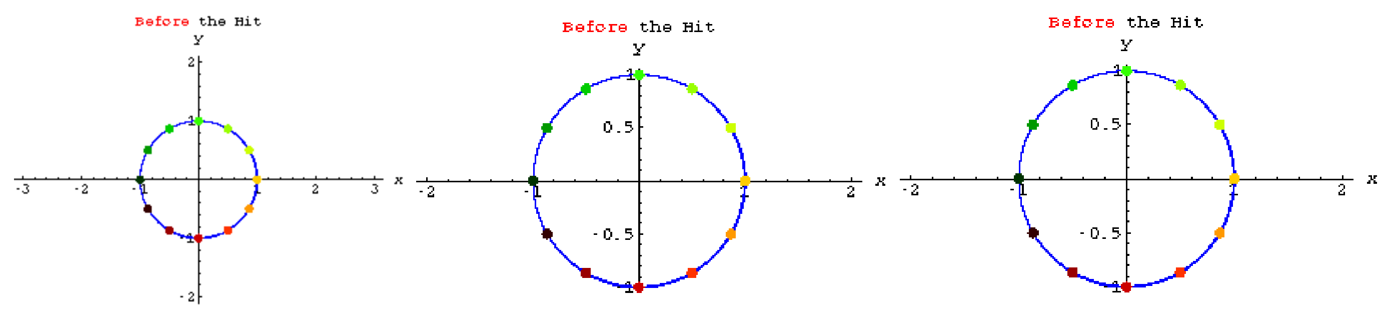


Matrix Stretchers

Defn The diagonal matrix $\left[\begin{smallmatrix} a & 0 \\ 0 & b \end{smallmatrix}\right]$ stretches in the x direction by a factor of a and in the y direction by a factor of b:

$$\begin{bmatrix} a & 0 \\ 0 & b \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} ax \\ by \end{bmatrix}$$

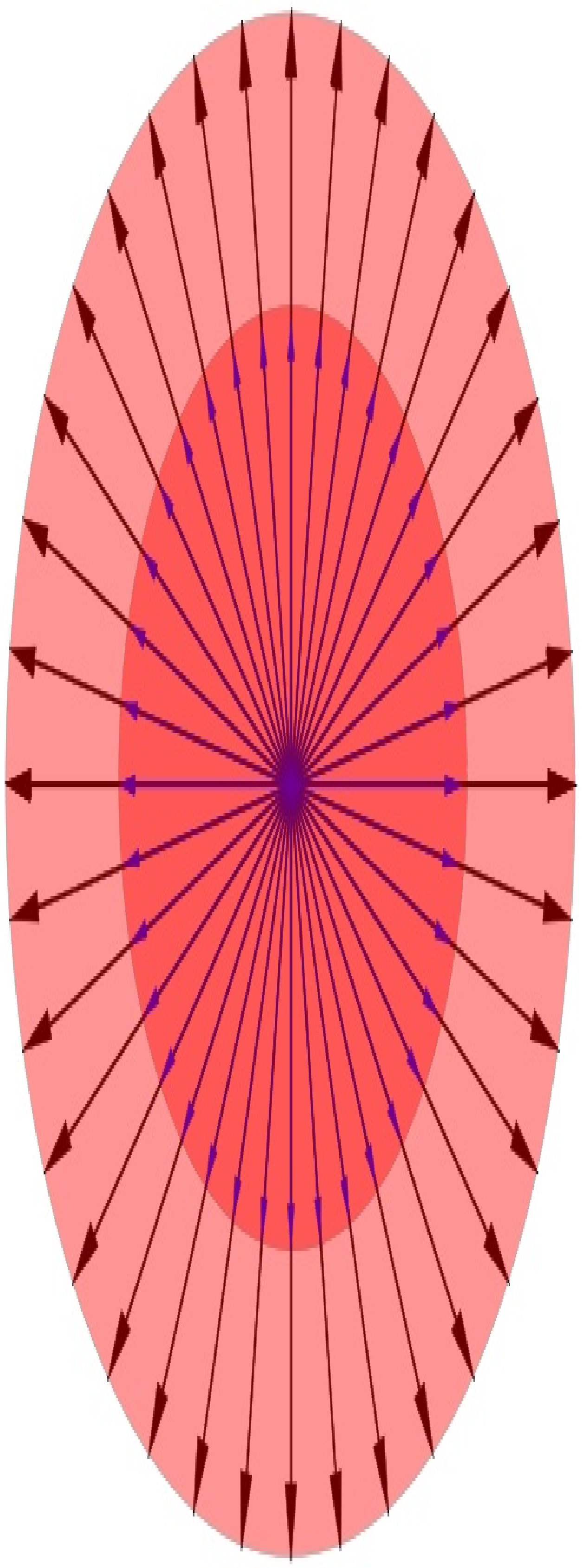
- What if we *stretch* a set of points with a matrix?
  - For example, with matrix $\left[\begin{smallmatrix} 3 & 0 \\ 0 & 2 \end{smallmatrix}\right]$, $\left[\begin{smallmatrix} 2 & 0 \\ 0 & 1/2 \end{smallmatrix}\right]$ and $\left[\begin{smallmatrix} 2 & 0 \\ 0 & 0 \end{smallmatrix}\right]$ to *stretch* a *unit circle* respectively.

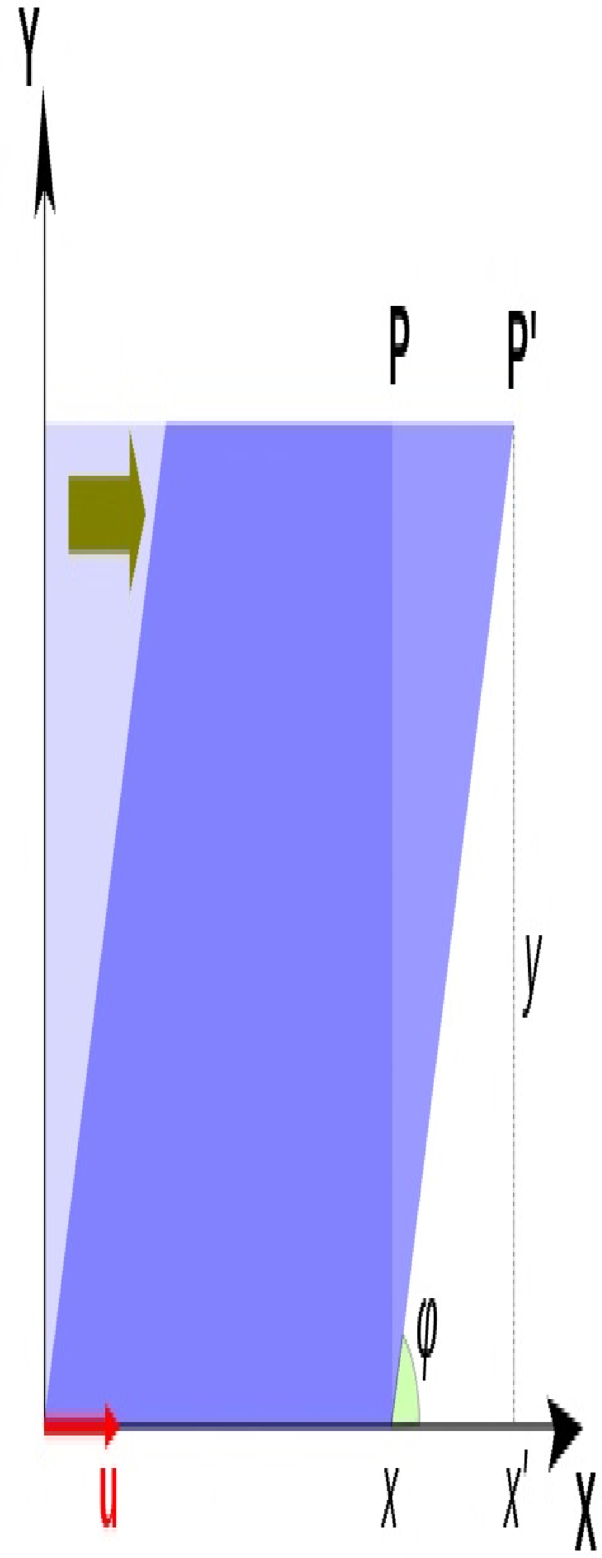


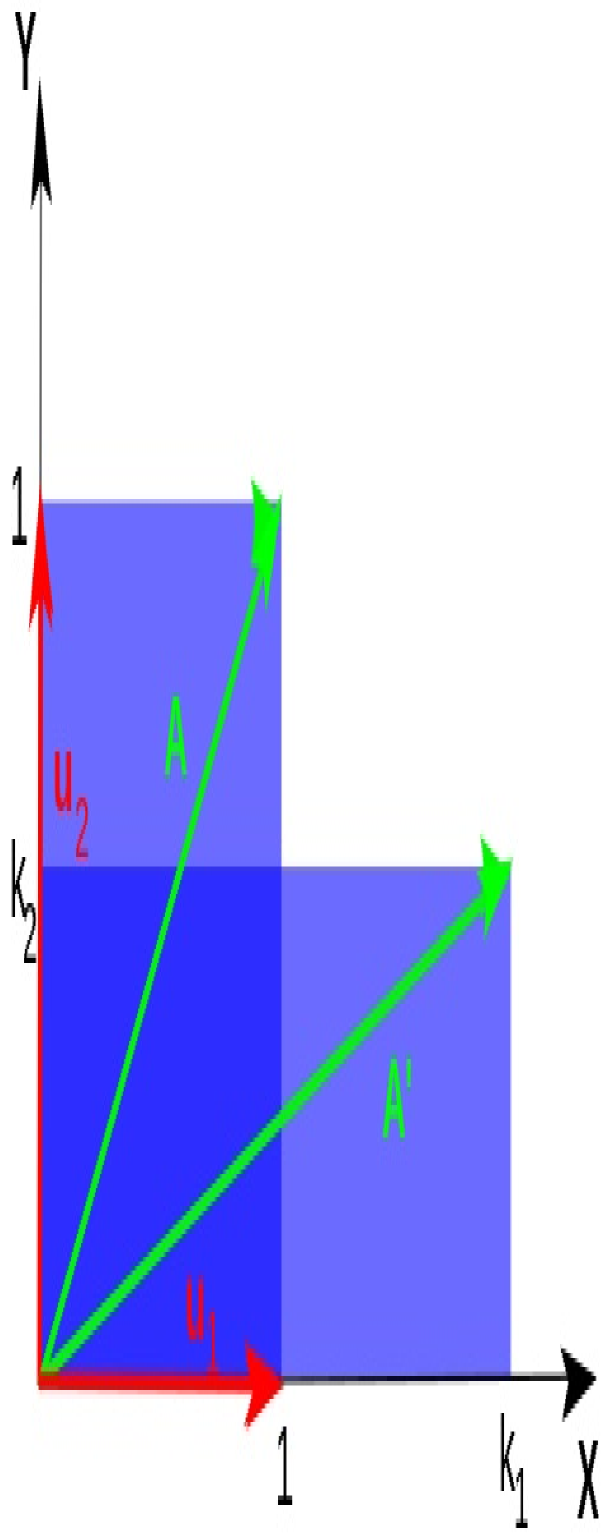
Exercises

Exercises:

- Define the matrix for the following transformation:







Matrix Permutation

Defn The *permutation matrix* P is from row shuffling of *identify matrix* I, and multiplying a matrix A by P exchanges the rows of A.

- A single *permutation matrix* P_{ij} exchanges row i and row j.

$$P_{12}A = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix} = \begin{bmatrix} d & e & f \\ a & b & c \\ g & h & i \end{bmatrix}$$

- A *permutation matrix* can do both of the row changes at once. It is the product of two separate single *permutation matrix* P_{23} and P_{12}.

$$P_{23}P_{12} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix}$$

- For any single *permutation matrix* P_{ij}, we have $P_{ij}^{-1} = P^T = P_{ij}$, namely $P_{ij}P_{ij} = I$
- Any combined permutation can also be reversed, e.g. $(P_{12}P_{23})(P_{23}P_{12}) = I$.

Matrix Elimination

Defn The *elementary matrix* P is changed from a single row in *identify matrix* I, and multiplying a matrix A by carefully designed P can eliminate a value in A.

- *Elementary matrix* E (subtract twice row1 from row2) and F (add row1 to row3).

$$EA = \begin{bmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix}, \qquad FA = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix} \begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix}$$

- Several *elementary matrices* can be multiplied together, and this generates a matrix that eliminate some values sequentially.

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

- Any *elementary matrix* E can be reversed by E^{-1}: $E^{-1}E = [\begin{smallmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ 0 & 0 & 1 \end{smallmatrix}][\begin{smallmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 0 & 0 & 1 \end{smallmatrix}] = I$.
- Hence a combined elimination GEF can be reversed by $F^{-1}E^{-1}G^{-1}$.
- All *elementary matrices*, *combinations* and *inverse* are *lower triangular* matrix.

Subspaces (1)

Defn A *subspace* of a *real space* is a non-empty subset that satisfies:

1. if any vectors x and y in the subspace, their sum $x+y$ is in the subspace;
2. if any vector x is multiplied by any scalar c, the multiple cx is in the subspace.

- The *column space* $\mathscr{C}(A)$ of a matrix $A \in \mathbb{R}^{m\times n}$ consists of all combinations of the columns.
  - vectors in $\mathscr{C}(A)$ are in $\mathbb{R}^m$
  - Ax is the combination of the columns, and $Ax \in \mathscr{C}(A)$.
  - $Ax = b$ can be solved if and only if b lies in the *column space* of A
- The *nullspace* $\mathscr{N}(A)$ of a matrix $A \in \mathbb{R}^{m\times n}$ consists of all vectors x such that $Ax = 0$.
  - vectors in $\mathscr{N}(A)$ are in $\mathbb{R}^n$

Subspaces (2)

Defn A *subspace* of a *real space* is a non-empty subset that satisfies:

1. if any vectors x and y in the subspace, their sum $x+y$ is in the subspace;
2. if any vector x is multiplied by any scalar c, the multiple cx is in the subspace.

- The *row space* $\mathscr{R}(A^T)$ of a matrix $A \in \mathbb{R}^{m\times n}$ consists of all combinations of the rows.
  - vectors in $\mathscr{R}(A^T)$ are in $\mathbb{R}^n$
  - yA is the combination of the rows, and $yA \in \mathscr{R}(A^T)$.
- The *left nullspace* $\mathscr{N}(A^T)$ of a matrix $A \in \mathbb{R}^{m\times n}$ consists of all vectors y such that $y^T A = 0$.
  - vectors in $\mathscr{N}(A^T)$ are in $\mathbb{R}^m$

Matrix Factorization

Model

Let $A \in \mathbb{R}^{m\times n}$, $x \in \mathbb{R}^n$ and $b \in \mathbb{R}^m$. The system is

$$Ax = b$$

- We wish to decompose the matrix A by writing it as a product of two or more matrices:

  $$A_{m\times n} = B_{m\times k} C_{k\times n}, \quad A_{m\times n} = B_{m\times k} C_{k\times r} D_{r\times n},$$

  - This is done in such a way that the right side of the equation yields some useful information or insight to the nature of the data matrix A, or is in other ways useful for solving the problem at hand.
- There are numerous useful *matrix decompositions*, also known as *matrix factorization*.

Gaussian Elimination

Echelon Form

Apply the carefully designed elimination E on both side of $Ax = b$, the system can be simplified into $Ux = c$. U is upper triangular, with pivots pivots not necessarily on the main diagonal, forms a *staircase* patten, or *echelon form* such as:

$$Ux = \begin{bmatrix} \ast\ast & \ast & \ast & \ast & \ast & \ast & \ast \\ 0 & \ast\ast & \ast & \ast & \ast & \ast & \ast \\ 0 & 0 & 0 & \ast\ast & \ast & \ast & \ast \\ 0 & 0 & 0 & 0 & 0 & 0 & \ast\ast \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} x = \begin{bmatrix} c_1 \\ c_2 \\ c_3 \\ c_4 \\ c_5 \end{bmatrix}$$

- For $Ux = c$, the unknowns in x go into two groups:

  ***basic* variables** correspond to *columns with pivots*
  ***free* variables** correspond to *columns without pivots*

- U can be further eliminated into the *reduced row echelon form* with:
  1. All non-zero rows are above any rows of all zeroes.
  2. Every leading coefficient is 1 and is the only non-zero entry in its column.
- the *rank* of A = the number of *pivots* in U
  1. r *basic variables* in matrix A, U
  2. $(n - r)$ *free variables* in matrix A, U

LU Factorization

LU

If no row exchanges required, m by n matrix A can be eliminated to an *upper triangular* matrix U using *Gaussian Elimination*: $EA = U$, where the m by m matrix E is a sequence of *elementary matrix*, with inverse $L = E^{-1}$. Hence we have $A = LU$.

- L is *lower triangular*, with 1s on its diagonal and l_{ij} below the diagonal;
- U is the *upper triangular* matrix its diagonal entries are the *pivots*.
- If row exchanges P are required, then we have $PA = LU$.

- The U can be further divided into a diagonal matrix D with the pivots, and then we can have $A = LDU$, where L and U have 1s on the diagonal, and D is the diagonal matrix of pivots.

$$U = \begin{bmatrix} d_1 & & & \\ & d_2 & & \\ & & . & \\ & & & d_n \end{bmatrix} \begin{bmatrix} 1 & u_{12}/d_1 & u_{13}/d_1 & \\ & 1 & u_{23}/d_2 & \\ & & . & \\ & & & 1 \end{bmatrix}$$

- The LDU factorization of a matrix A is unique.
- If A is symmetric, and if it can be factored without row exchange by permutation matrix P, then $A = LDL^T$.

LU Factorization

Exercises:

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

  1. Compute the LU factorization
  2. Determine a set of basic variables and a set of free variables
  3. What is the rank of A?
  4. Find the general solution to $Ax = 0$. (do it later)

Solution of $Ax = b$

$$Ux = \begin{bmatrix} * & * & * & * & * & * & * \\ 0 & * & * & * & * & * & * \\ 0 & 0 & 0 & * & * & * & * \\ 0 & 0 & 0 & 0 & 0 & 0 & * \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} x = \begin{bmatrix} c_1 \\ c_2 \\ c_3 \\ c_4 \\ c_5 \end{bmatrix}$$

- When $Ax = b$ (or $Ux = c$) has solution(s)?
  - b lies in the *column space* of A
  - or the last $(m - r)$ elements of c are 0
- A solution $x_{particular}$ to $Ux = c$?
  1. Set all $(n - r)$ free variables to 0
  2. For the r variables, from r rows, we find a particular solution $x_{particular}$
- If a solution $x_{homogeneous}$ makes $Ux = 0$, then solutions to $Ax = b$ (or $Ux = c$) are

$$x_{general} = x_{homogeneous} + x_{particular} \qquad x_{homogeneous} \in \mathcal{N}(A)$$

- When the rank r is as large as possible, then

  [$r = n$] no free variables in x, the *nullspace* $\mathcal{N}(A)$ contains only $x = 0$

  [$r = m$] no constraint on c (or b), there is always a solution x.

Solution of $Ax = 0$

$$URx = \begin{bmatrix} *1 & *0 & * & *0 & * & * & *0 \\ 0 & *1 & * & *0 & * & * & *0 \\ 0 & 0 & 0 & *1 & * & * & *0 \\ 0 & 0 & 0 & 0 & 0 & 0 & *1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} x = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} \Longrightarrow \begin{bmatrix} I_{rr} & F_{r(n-r)} \\ 0_{(m-r)r} & 0_{(m-r)(n-r)} \end{bmatrix} y = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

- We further reduce U into the *reduced row echelon form*.
- General Case: shuffling x, $Ux = 0$ changes into $Ry = 0$, and we have $y = \binom{-F}{I}$.
- For example:

$$\begin{bmatrix} 1 & 2 & 3 \\ 2 & 4 & 6 \\ 2 & 6 & 8 \\ 2 & 8 & 10 \end{bmatrix} \Longrightarrow \begin{bmatrix} 1 & 2 & 3 \\ 0 & 2 & 2 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \Longrightarrow \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \quad \text{Solution:} y = \begin{bmatrix} -F \\ I \end{bmatrix} = \begin{bmatrix} -1 \\ -1 \\ 1 \end{bmatrix}$$

- Special Case (1): $r = n < m$

$$Rx = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix} x = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

  - there is only one trivial solution to $Ax = 0$: *zero vectors*.
  - $R = \binom{I_{rr}}{0}$, and the *nullspace* $\mathcal{N}(A)$ only contains zero vector.
  - There is only one solution to $Ax = b$
- Special Case (2): $r = m < n$

$$Rx = \begin{bmatrix} 1 & 0 & 0 & * \\ 0 & 1 & 0 & * \\ 0 & 0 & 1 & * \end{bmatrix} x = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

  - there is always one or more solutions to $Ax = 0$: $y = \binom{-F}{I}$..
  - $R = (I_{rr}\ F_{r(n-r)})$, solution $x_{particular}$ always exists for $Ax = b$.
  - There is one or infinity solutions to $Ax = b$
- Special Case (3): $r = m = n$

$$Rx = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} x = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

  - there is always one solution to $Ax = 0$: *zero vectors*.
  - $R = I_{rr}$, the *nullspace* $\mathcal{N}(A)$ only contains zero vector.
  - There is only one solution $x_{particular}$ to $Ax = b$
- Special Cases:

  [$r = n < m$] [a] we have $R = \binom{I_{rr}}{0}$, and the *nullspace* $\mathcal{N}(A)$ only contains zero vector. there is zero or one solution to $Ax = b$.

  [$r = m < n$] [b] we have $R = (I_{rr}\ F_{r(n-r)})$, one or more solutions $x_{particular}$ always exists for $Ax = b$.

  [$r = m = n$] we have $R = I_{rr}$, the *nullspace* $\mathcal{N}(A)$ only contains zero vector, there is only one solution $x_{particular}$ to $Ax = b$.

[a] the matrix is *Full Column Rank*

[b] the matrix is *Full Row Rank*

Solution of $Ax = b$

Exercises:

- $Ux = \begin{bmatrix} 1 & 2 & 3 & 4 \\ 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x4 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$

  1. Find all solutions to the above equation
  2. If the right side is changed from $(0,0,0)$ to $(a,b,0)$, what are the solutions?

Solution of $Ax = b$

Exercises:

- $Ax = \begin{bmatrix} 2 \\ 4 \\ 2 \end{bmatrix} \qquad x = \left[\begin{smallmatrix} 2 \\ 0 \\ 0 \end{smallmatrix}\right] + c \left[\begin{smallmatrix} 1 \\ 1 \\ 0 \end{smallmatrix}\right] + d \left[\begin{smallmatrix} 0 \\ 0 \\ 1 \end{smallmatrix}\right]$

  1. What is the dimension of $\mathcal{N}(A)$
  2. What is A?
  3. When $Ax = b$ can be solved?

Linear Independence

Defn Given a set of vectors $v_1, v_2, \cdots, v_k$ in $\mathbb{R}^m$, consider the set of linear combinations $y = \sum_{i=1}^{k} c_i v_i$ for arbitrary coefficients c_i. The vectors $(v_i)_{i=1}^{k}$ are *linear independent*, if $\sum_{i=1}^{k} c_i v_i = 0$ if and only if $c_i = 0$ for all $i = 1, \cdots, k$

- A set of m linear independent vectors of $\mathbb{R}^m$ is called a *basis* in $\mathbb{R}^m$.
- Any vector in $\mathbb{R}^m$ can be expressed as a linear combination of the *basis* vectors.
- Example: The columns of the matrix

$$\begin{bmatrix} v_1 & v_2 & v_3 & v_4 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 1 \\ 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 1 \end{bmatrix}$$

  are not linear independent, as we have $c_1 v_1 + c_2 v_2 + c_3 v_3 + c_4 v_4 = 0$ holds for $c_1 = c_3 = 1$, $c_2 = c_4 = -1$.
- The columns of the matrix are *linear independent*, is the same as the *nullspace* of the matrix contains only the *zero vector*.

Dimension of a Space

Defn If a vector space V consists of all linear combinations of the particular vectors $v_1, v_2, \cdots, v_k$, then these vectors span the space. Every vector v in V can be expressed as a combination of v_i. A basis for a vector space is a set of vectors having two properties at once:

1. It is linearly independent
2. It spans the space

- Any two bases for a vector space V contain the same number of vectors. This number is called the dimension of V.
- In a subspace of dimension k, no set of more than k vectors can be linearly independent; no set of fewer than k vectors can span the space.
- Exercises:
  - find a basis for the *nullspace* of following matrix

$$A = \begin{bmatrix} 0 & 1 & 2 & 3 & 4 \\ 0 & 1 & 2 & 4 & 6 \\ 0 & 0 & 0 & 1 & 2 \end{bmatrix} = LU = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 0 & 1 & 1 \end{bmatrix} \begin{bmatrix} 0 & 1 & 2 & 3 & 4 \\ 0 & 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

  - find the general solution to above matrix $Ax = \begin{bmatrix} 3 \\ 4 \\ 1 \end{bmatrix}$.

Rank of a Matrix

Rank The *rank* of a matrix is the number of *pivots* in its *reduced row echelon form*, which is equal to the number of *genuinely independent rows* in the matrix.

- it is also the number of *genuinely independent columns* in the matrix.
- A square matrix $A \in \mathbb{R}^{n \times n}$ with rank n is called *nonsingular*, and it has an inverse A^{-1} satisfying $AA^{-1} = A^{-1}A = I$
- Example: The rank of the following matrix is 1

$$A = \begin{bmatrix} 2 & 1 & 1 \\ 4 & 2 & 2 \\ 8 & 4 & 4 \\ -2 & -1 & -1 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \\ 4 \\ -1 \end{bmatrix} \begin{bmatrix} 2 & 1 & 1 \end{bmatrix}$$

- Every matrix of rank one has the simple form $A = uv^T = \begin{bmatrix} v_1 u & v_2 u & \cdots & v_n u \end{bmatrix}$, with all columns (and all rows) are linear dependent.
- Exercise:
  - Rewrite the following matrix $A = \begin{bmatrix} 1 & 0 & 0 & 3 \\ 0 & 0 & 0 & 0 \\ 2 & 0 & 0 & 6 \end{bmatrix}$.

Questions?

Contact Information

Associate Professor **Gang Li**
School of Information Technology
Deakin University, Australia

gangli@tulip.org.au

Team for Universal Learning and Intelligent Processing