Chapter 2: Inner Product Spaces

- Motivation
- Definition of inner product
- The spaces L^2 and l^2
- Schwarz and triangular inequalities
- Orthogonality
- Linear operators
- Least squares

1. Motivation

- For two vectors X=(x_1,x_2,x_3), Y=(y_1,y_2,y_3) in R^3, the standard inner product of X and Y is defined as

$$\langle X, Y \rangle = x_1 y_1 + x_2 y_2 + x_3 y_3$$

- This definition is partly motivated by the desire to measure the length of a vector (Pythagorean theorem)

$$\text{Length of } X = \sqrt{x_1^2 + x_2^2 + x_3^2} = \sqrt{\langle X, X \rangle}$$

- If X is a unit vector, then $\langle X, Y \rangle$ is to measure the vector length of Y, i.e., the projection of Y on vector X.

- Ex.點(3,6,5)到點(3,8,7)的歐氏距離為：

$$\sqrt{(3-3)^2+(8-6)^2+(7-5)^2}=\sqrt{0+4+4}=\sqrt{8}$$

- Ex.向量(3,6,5)長度為：

$$\sqrt{3^2+6^2+5^2}=\sqrt{70}$$

- Ex.如果X不是單位向量(ex:(1,0,0))，則<X,Y>不再是量測Y投影到X方向的長度，而是有放大或縮小的效果 : ex: X=(2,0,0) or X=(0.5,0,0)

2 Definition of inner product

- For any dimension n, the two vectors

$$X = (x_1, x_2, \cdots, x_n) \text{ and } Y = (y_1, y_2, \cdots, y_n) \text{ ,}$$

the Euclidean inner product is

$$\langle X, Y \rangle = \sum_{j=1}^{n} x_j y_j$$

- Complex form: Z and W are both complex vectors

$$\langle Z, W \rangle = \sum_{j=1}^{n} z_j \overline{w}_j$$

conjugate

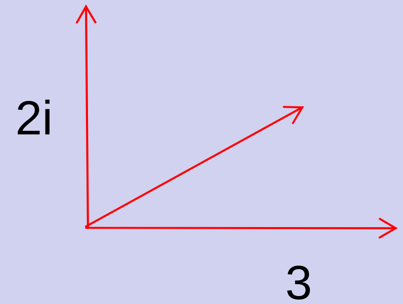
- 複數的向量內積使用共軛複數的原因
- If X=3+2i=(3，2i)， then the length of X is：

Without conjugate (wrong):

$$\sqrt{< X, X >} = \sqrt{(3,2i)\cdot(3,2i)} = \sqrt{9+4i^2} = \sqrt{9-4} = \sqrt{5}$$

With conjugate (correct):

$$\sqrt{< X, X >} = \sqrt{(3,2i)\cdot(3,-2i)} = \sqrt{9-4i^2} = \sqrt{9+4} = \sqrt{13}$$



3 The spaces L^2 and l^2 (1)

- Continuous form (連續型): L^2

The energy of a continuous function *f* in the interval is defined as:

$$L^2([a,b]) = \left\{ f : [a,b] \rightarrow C; \int_a^b | f(t) |^2 \, dt < \infty \right\}$$

The L^2 inner product in the interval [a, b] of two continuous functions is defined as:

$$\langle f, g \rangle_{L^2} = \int_a^b f(t) \overline{g}(t) dt$$

連續型內積定義

(f and g can be complex.)

能量的定義

- 能量的定義，就是將每一個element平方後相加，也就是向量長度的平方。
- Ex. The energy of a vector (or a signal) X=(3,4,5) is defined by: $3^2+4^2+5^2=50$
- Ex. The energy of a vector (or a signal)

  X=(x1,x2,x3,…,xn) is defined by:

$$<X,X>=x_1^2+x_2^2+x_3^2+\cdots+x_n^2$$

3 The spaces L^2 and l^2 (2)

- Discrete form (離散型): l^2

The l^2 inner product in the interval [a, b] of two discrete functions is defined as:

$$\left\langle X, Y \right\rangle_{l^2} = \sum_{i=a}^{b} x_i \overline{y}_i$$

離散型內積定義

(X and Y can be complex.)

4 Schwarz inequalities

Schwarz inequality:

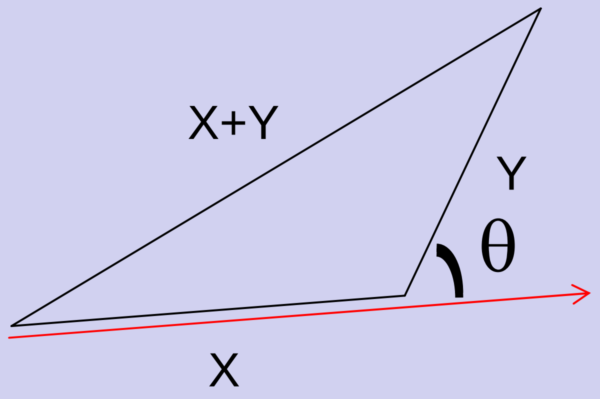
$$|\langle X, Y \rangle| \leq \| X \| \| Y \|$$

Equality holds if and only if X and Y are linear dependent.
If X and Y are linear independent, what happens?

$$|\langle X, Y \rangle| = \| X \| \| Y \| \cos\theta \leq \| X \| \| Y \|$$

線性相依

$$\text{If } \theta = 0 \text{ or } \theta = 180°, \cos\theta = 1.$$



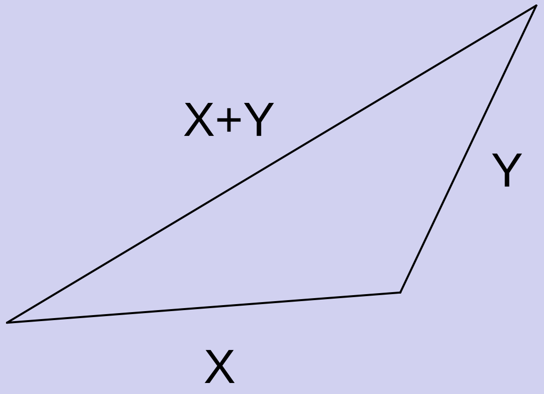
線性無關

$$\text{If } \theta = 90° \text{ or } \theta = 270°, \cos\theta = 0.$$

Triangular inequalities

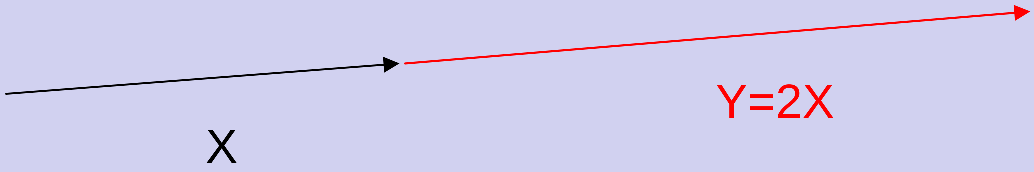
Triangle inequality:

$$\| X + Y \| \le \| X \| + \| Y \|$$



Equality holds if and only if X or Y is a <u>positive multiple</u> of the other.

$$X = tY, \quad t > 0$$



$$\| X + Y \| = \| X \| + \| Y \|$$

5 Orthogonality 正交性 (1)

The vectors X and Y are said to be orthogonal if

$$\langle X, Y \rangle = 0$$

內積為0

The vectors X and Y are said to be orthonormal if

$$\begin{cases} \langle X, Y \rangle = 0 \\ \| X \| = 1 \text{ and } \| Y \| = 1 \end{cases}$$

內積為0

長度為1

5 Orthogonality (2)

Example: The function $f(t) = \sin t$ and $g(t) = \cos t$ are orthogonal in $L^2([-\pi, \pi])$.

Proof:

$$\langle f, g \rangle = \int_{-\pi}^{\pi} \sin t \cos t dt$$

$$= \frac{1}{2} \int_{-\pi}^{\pi} \sin(2t) dt$$

$$= \frac{-1}{4} \cos(2t) |_{-\pi}^{\pi}$$

$$= 0$$

內積為0，所以正交(orthogonal)
是否為orthonormal?

5 Orthogonality (3)

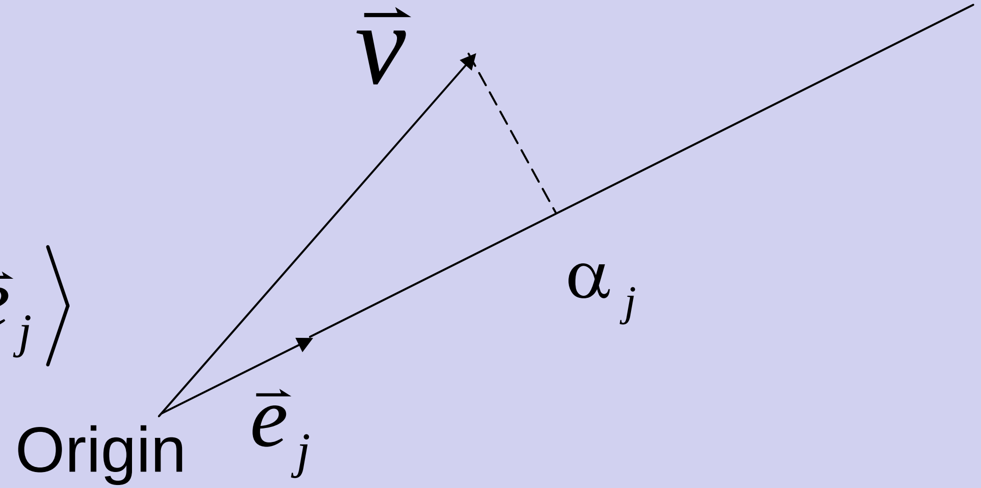
- Orthogonal projection

$\{\vec{e}_1, \vec{e}_2, \cdots, \vec{e}_N\}$ is an orthonormal collection of vectors

If $\vec{v}$ lies in the space of $\{\vec{e}_1, \vec{e}_2, ..., \vec{e}_N\}$, then,

$$\vec{v} = \sum_{j=1}^{N} \alpha_j \vec{e}_j$$

where $\alpha_j = \left\langle \vec{v}, \vec{e}_j \right\rangle$



That means: vector $\vec{v}$ can be spanned in a new coordinate system $\{\vec{e}_1, \vec{e}_2, ..., \vec{e}_N\}$ and its coordinate is $(\alpha_1, \alpha_2, ..., \alpha_N)$.

5 Orthogonality (4)

Example:

A vector in R^3 is $v = (3,5,7)$, it means

$$3 = \langle \vec{v}, \vec{x} \rangle \quad \text{where } \vec{x} = (1,0,0);$$

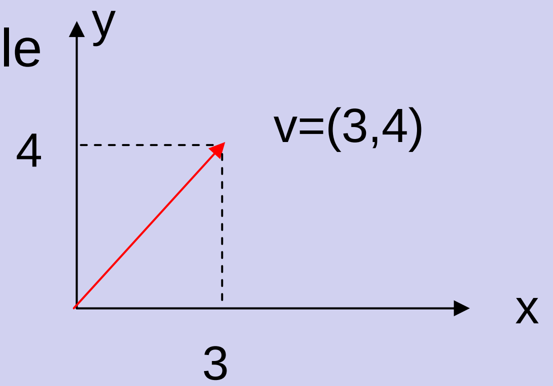
$$5 = \langle \vec{v}, \vec{y} \rangle \quad \text{where } \vec{y} = (0,1,0); \text{ and}$$

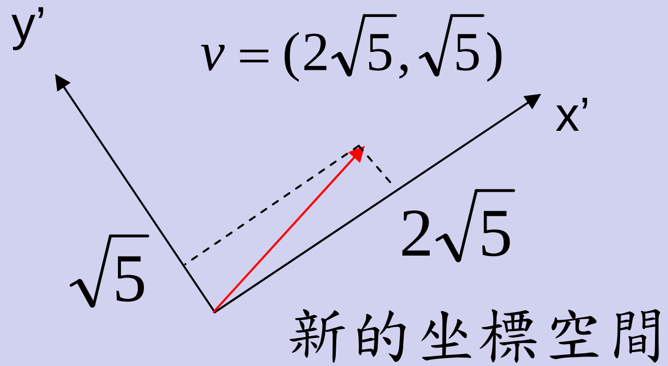
$$7 = \langle \vec{v}, \vec{z} \rangle \quad \text{where } \vec{z} = (0,0,1).$$

$$\Rightarrow \vec{v} = 3\vec{x} + 5\vec{y} + 7\vec{z}$$

任何向量都可以投射到新的坐標空間

Example





$$\begin{cases} x' = (2,1)/\sqrt{5} \\ y' = (-1,2)/\sqrt{5} \end{cases}$$

$$(3,4)\begin{pmatrix} 2 \\ 1 \end{pmatrix}/\sqrt{5} = 10/\sqrt{5} = 2\sqrt{5} = 4.4721$$

$$(3,4)\begin{pmatrix} -1 \\ 2 \end{pmatrix}/\sqrt{5} = 5/\sqrt{5} = \sqrt{5}$$

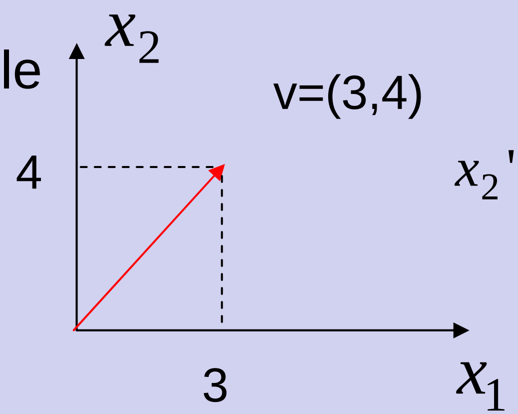
$$\text{energy} = 3^2 + 4^2 = 9 + 16 = 25$$ **(x,y)-plane**

$$\text{energy} = (2\sqrt{5})^2 + (\sqrt{5})^2 = ?$$ **(x',y')-plane**

投射後能量守恆

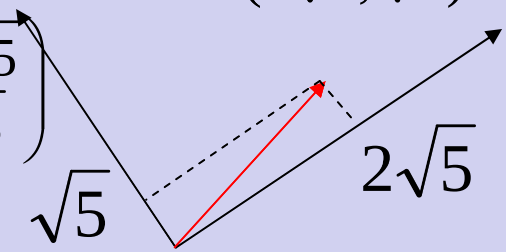
矩陣表示　向量在投射後的坐標，可用矩陣表示法計算

Example



$$v=(2\sqrt{5},\sqrt{5})$$

$$x_2'=\begin{pmatrix}-1/\sqrt{5}\\ 2/\sqrt{5}\end{pmatrix}\qquad x_1'=\begin{pmatrix}2/\sqrt{5}\\ 1/\sqrt{5}\end{pmatrix}$$



新的坐標空間

$$\begin{pmatrix}3\\4\end{pmatrix}=\begin{pmatrix}\vec{x}_1^T\\ \vec{x}_2^T\end{pmatrix}\vec{v}=\begin{pmatrix}1&0\\0&1\end{pmatrix}\begin{pmatrix}3\\4\end{pmatrix}\qquad \begin{pmatrix}2\sqrt{5}\\ \sqrt{5}\end{pmatrix}=\begin{pmatrix}\vec{x}'^T_1\\ \vec{x}'^T_2\end{pmatrix}\vec{v}=\begin{pmatrix}2/\sqrt{5}&1/\sqrt{5}\\-1/\sqrt{5}&2/\sqrt{5}\end{pmatrix}\begin{pmatrix}3\\4\end{pmatrix}$$

本位座標空間　新的坐標　新的坐標空間

用途：在找出多數向量的主軸(主軸分析)，資料壓縮

Example: 建立新的坐標系統，使得Z軸與給定的向量平行

Let $\vec{v} = (3,5,7)$ be a vector in Cartison coordinate. Construct a new coordinate system, in which $\dfrac{\vec{v}}{\|\vec{v}\|}$ is a new $\vec{z}'$ axis.

$$\text{Define } \vec{e}_1 = \frac{1}{\sqrt{14}}(3,1,-2) \text{ and } \vec{e}_3 = \frac{\vec{v}}{\|\vec{v}\|} = \frac{1}{\sqrt{83}}(3,5,7),$$

so $\langle \vec{e}_1, \vec{e}_3 \rangle = 0$

Define $\vec{e}_2 = (x, y, z)$ which satisfies: $\langle e_1, e_2 \rangle = 0$ and $\langle e_3, e_2 \rangle = 0$.

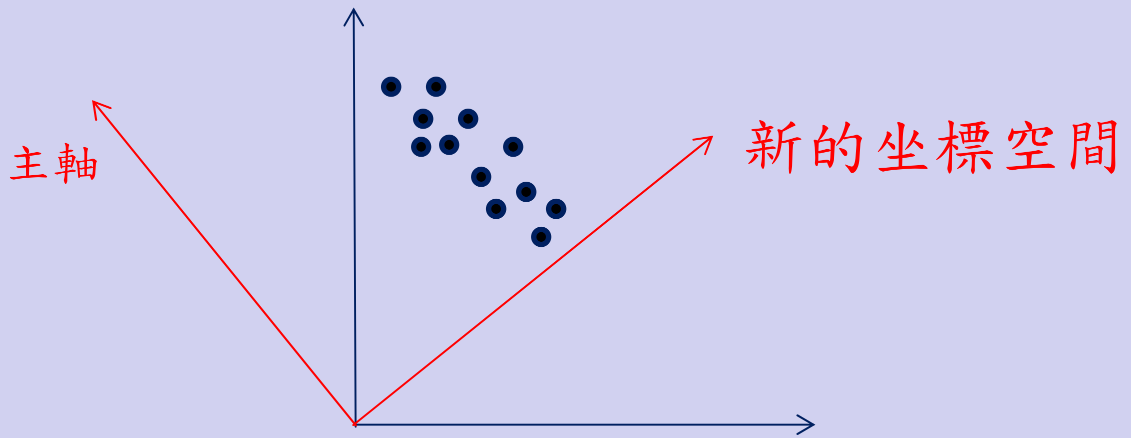
$$\begin{cases} 3x + y - 2z = 0 \\ 3x + 5y + 7z = 0 \end{cases} \Rightarrow y = \frac{-9}{4}z,\ x = \frac{17}{12}z.$$

$$\Rightarrow e_2 = \left(\frac{17}{12}, \frac{-9}{4}, 1\right)\frac{1}{\sqrt{8.0694}}$$

Therefore, in the new coordinate system, $\vec{v}$ can be represented as: $0\vec{e}_1 + 0\vec{e}_2 + \sqrt{83}\vec{e}_3$

or in a matrix form: $(\vec{e}_1, \vec{e}_2, \vec{e}_3)\begin{pmatrix} 0 \\ 0 \\ \sqrt{83} \end{pmatrix}$

找出多數向量的主軸(主軸分析)



找出主軸後，次要的軸就變得不重要，可以忽略，因此可以做：
1.資料壓縮 2.降低維度

任何一個一維空間的資料，可視為超高維空間的一個點(或點向量)

$$\text{1D data}: f = \left(f_1, f_2, \ldots, f_n\right) \quad \text{length} = n$$

$$\text{In } n-\text{D}, \vec{f} = \left(f_1, f_2, \ldots, f_n\right) \Leftrightarrow f_1\begin{pmatrix}1\\0\\\vdots\\0\end{pmatrix} + f_2\begin{pmatrix}0\\1\\0\\\vdots\end{pmatrix} + \ldots + f_n\begin{pmatrix}0\\\vdots\\0\\1\end{pmatrix}$$

If there is another coordinate system: $\{\vec{e}_1, \vec{e}_2, \cdots, \vec{e}_n\} \in R^n$

then f can be represented in a new form:

$$\vec{f} = \left\langle \vec{f}_1, \vec{e}_1 \right\rangle \vec{e}_1 + \left\langle \vec{f}_1, \vec{e}_2 \right\rangle \vec{e}_2 + \ldots + \left\langle \vec{f}_1, \vec{e}_n \right\rangle \vec{e}_n$$

此一點向量，可以投射到任何一個新的(維度相同的)座標系統，這是Fourier transform的理論基礎。

$$T(\vec{f}) = \vec{g} = \begin{pmatrix}\vec{e}_1^{\,T}\\\vec{e}_2^{\,T}\\\vdots\\\vec{e}_n^{\,T}\end{pmatrix}_{n\times n \text{ matrix}} \vec{f}$$

$$\vec{g} = (\left\langle \vec{f}, \vec{e}_1 \right\rangle, \left\langle \vec{f}, \vec{e}_2 \right\rangle, \ldots, \left\langle \vec{f}, \vec{e}_n \right\rangle)$$

6 Linear operators (1)

Definition:

A linear operator (or map) between a vector space V and a vector space W is a function $\mathrm{T}: \mathrm{V} \rightarrow \mathrm{W}$ which satisfies $T(au + bv) = aT(u) + bT(v)$ for $u, v \in \mathrm{V}$ $a, b \in C$.

If V and W are finite dimensional, then T can be represented as a matrix. For any vector $v = \sum_j x_j \vec{v}_j$:

$$T(\vec{v}) = \sum_i \sum_j t_{ij} x_j = \begin{pmatrix} t_{11} & \cdots & t_{1n} \\ \vdots & \ddots & \vdots \\ t_{m1} & \cdots & t_{mn} \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$$

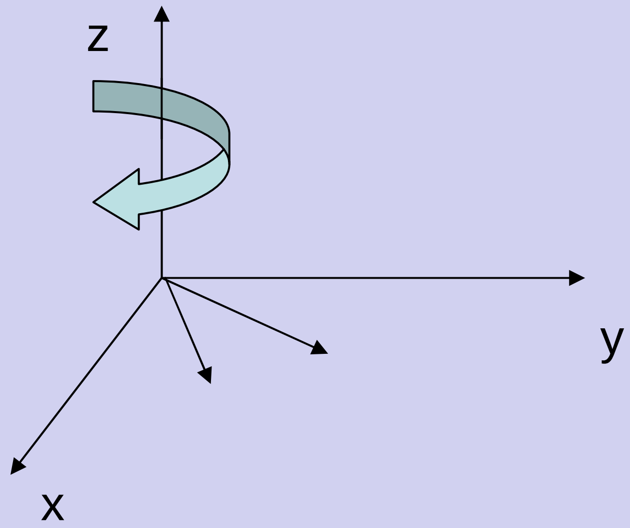
and $T(\vec{v}_j) = \sum_{i=1}^{m} t_{ij} \vec{w}_j$

用途：空間位置的轉換，座標系統轉換等。

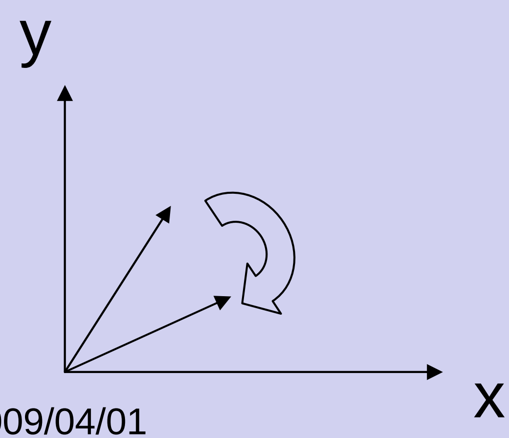
6 Linear operators (2)

Example

Rotate a vector $v = (1,2)$ with respect to z - axis for an angle θ.



$$T(v) = \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} 1 \\ 2 \end{pmatrix}$$



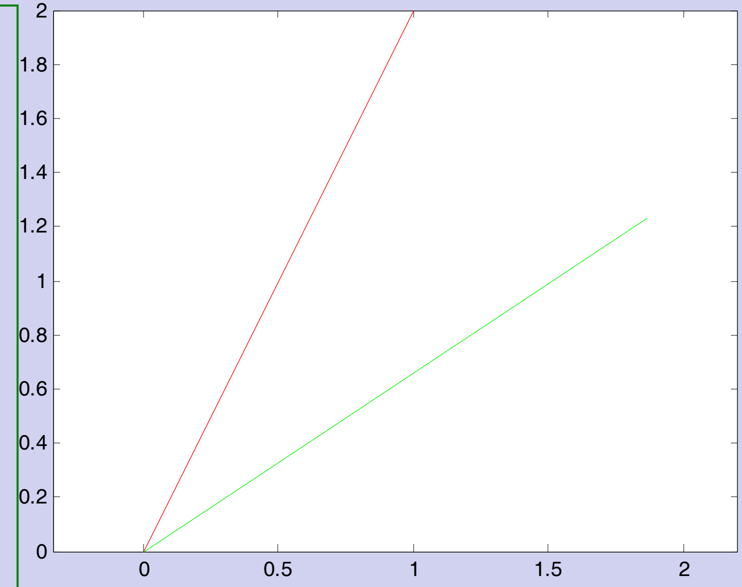
6 Linear operators (3)

Matlab program: 旋轉30度

```
v=[1;2];
a=30/180*pi;
t=[cos(a) sin(a);-sin(a) cos(a)];

w=t*v;
figure(1); clf;
plot([0 v(1)],[0 v(2)],'r');
hold on;
plot([0 w(1)],[0 w(2)],'g');
axis equal;
```

Result:



Red: Before rotation,
Green: After rotation.

6 Linear operators (4)

Adjoints: Definition 既然可以轉過去，當然也可以轉回來

If $T:V\rightarrow W$ is a linear operator between two inner product spaces, the adjoint of T is the linear operator $T^{*}:W\rightarrow V$ that satisfies

$$\left\langle T(v),w\right\rangle \mathrm{w}=\left\langle v,T^{*}(w)\right\rangle \mathrm{v}$$

Let $\vec{v}=\sum_{j=1}^{n}x_j\vec{v}_j$ $\quad\because\ T(\vec{v}_j)=\sum_{i=1}^{m}a_{ij}\vec{w}_i$

$$\therefore\ T(\vec{v})=T(\sum_{j=1}^{n}x_j\vec{v}_j)=\sum_{j=1}^{n}x_jT(\vec{v}_j)=\sum_{i=1}^{m}\sum_{j=1}^{n}(a_{ij}x_j)\vec{w}_i$$

$$=\sum_{i=1}^{m}c_i\vec{w}_i=\vec{w}\quad\text{where}\quad c_i=\sum_{j=1}^{n}a_{ij}x_j$$

If $T^{*}(\vec{w})=\vec{v}$, then let $T^{*}(\vec{w}_i)=\sum_{j=1}^{n}b_{ij}\vec{v}_j$

6 Linear operators (5)

$$T^*(\vec{w}) = T^*\left(\sum_{i=1}^{m} c_i \vec{w}_i\right) = \sum_{i=1}^{m} c_i T^*(\vec{w}_i) = \sum_{i=1}^{m}\sum_{j=1}^{n} c_i b_{ij} \vec{v}_j$$

$$= \sum_{j=1}^{n}\sum_{i=1}^{m} c_i b_{ij} \vec{v}_j = \vec{v}$$

Known $\vec{v} = \sum_{j=1}^{n} x_j \vec{v}_j$ $\Rightarrow$ $\sum_{i=1}^{m} c_i b_{ij} = x_j$

Known $c_i = \sum_{j=1}^{n} a_{ij} x_j$ $\Rightarrow$ $x_j = \sum_{i=1}^{m}\sum_{j=1}^{n} a_{ij} b_{ij} x_j$

6 Linear operators (6)

Matrix representation:

Let $\vec{v} = (\vec{v}_1, \cdots, \vec{v}_n)\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ and $T(\vec{v}_j) = (\vec{w}_1, \cdots, \vec{w}_m)\begin{pmatrix} a_{1j} \\ \vdots \\ a_{mj} \end{pmatrix}$

$$\Rightarrow \quad T(\vec{v}) = T\left((\vec{v}_1, \cdots, \vec{v}_n)\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}\right) = T((\vec{v}_1, \cdots, \vec{v}_n))\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$$

$$= (\vec{w}_1, \cdots, \vec{w}_n)\begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix}\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$$

$$\Rightarrow \quad T(\vec{v}) = (\vec{w}_1, \cdots, \vec{w}_m)\begin{pmatrix} c_1 \\ \vdots \\ c_m \end{pmatrix} = \vec{w} \quad \Rightarrow \quad \mathbf{C=AX}$$

6 Linear operators (7)

Similarly

$$\vec{w} = (\vec{w}_1, \cdots, \vec{w}_m)\begin{pmatrix} c_1 \\ \vdots \\ c_m \end{pmatrix}$$ (known) Let $$T^*(\vec{w}_i) = (\vec{v}_1, \cdots, \vec{v}_n)\begin{pmatrix} b_{1i} \\ \vdots \\ b_{ni} \end{pmatrix}$$

$$\Rightarrow \quad T^*(\vec{w}) = (\vec{v}_1, \cdots, \vec{v}_n)\begin{pmatrix} b_{11} & \cdots & b_{1m} \\ \vdots & \ddots & \vdots \\ b_{n1} & \cdots & b_{nm} \end{pmatrix}\begin{pmatrix} c_1 \\ \vdots \\ c_m \end{pmatrix}$$

$$= \vec{v} = (\vec{v}_1, \cdots, \vec{v}_n)\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$$ $\Rightarrow$ **X=BC**

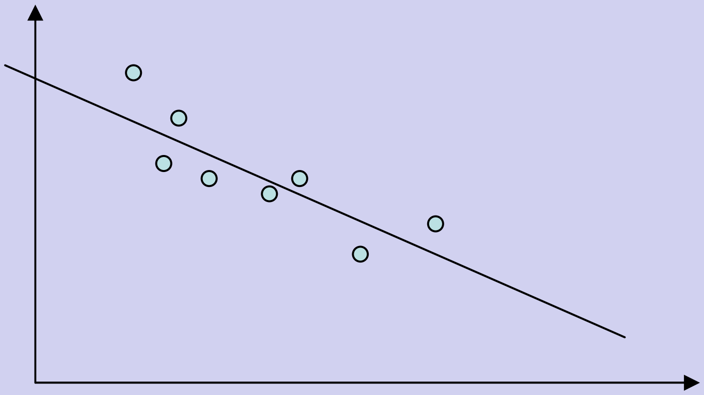
$\Rightarrow$ **X=BC=BAX** $\Rightarrow$ **BA=I**

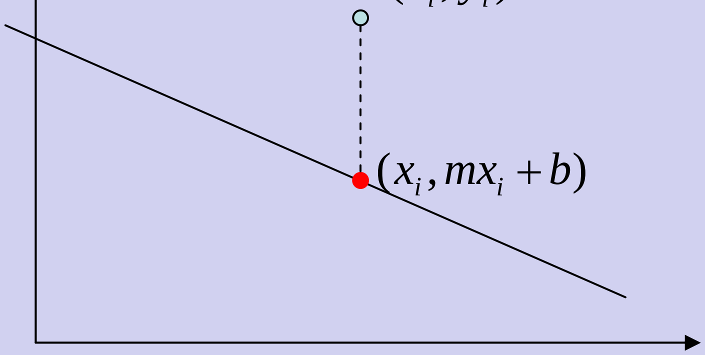
7 Least squares (1)

- Motivation:
  - 1. we often use least squares to develop a procedure for approximating signals (or functions).
  - 2. we often use least squares to get model parameters in a fitting problem.

7 Least squares (2)

Best line fitting problem (overdetermined problem)





Error at xi is $| y_i - (mx_i + b) |$

The best line fitting is to find the minimum total square error:

$$E = \sum_{i=1}^{N} (y_i - (mx_i + b))^2 \qquad N >> 2$$

7 Least squares (3)

$$E = \sum_{i=1}^{N} (y_i - (mx_i + b))^2$$

$$\begin{cases} \dfrac{\partial E}{\partial m} = 0 \quad \to \sum_{i=1}^{N} 2(y_i - (mx_i + b))(-x_i) = 0 \\ \dfrac{\partial E}{\partial b} = 0 \quad \to \sum_{i=1}^{N} 2(y_i - (mx_i + b))(-1) = 0 \end{cases}$$

$$\begin{cases} \sum_{i=1}^{N} (y_i x_i - m x_i^2 - b x_i) = 0 \\ \sum_{i=1}^{N} (y_i - m x_i - b) = 0 \end{cases}$$

$$\begin{cases} \sum_{i=1}^{N} y_i x_i = \sum_{i=1}^{N} m x_i^2 + \sum_{i=1}^{N} b x_i \\ \sum_{i=1}^{N} y_i \quad = \sum_{i=1}^{N} m x_i + \sum_{i=1}^{N} b \end{cases}$$

⇨ Linear equations.

$$\begin{pmatrix} \sum_{i=1}^{N} y_i x_i \\ \sum_{i=1}^{N} y_i \end{pmatrix} = \begin{pmatrix} \sum_{i=1}^{N} x_i^2 & \sum_{i=1}^{N} x_i \\ \sum_{i=1}^{N} x_i & N \end{pmatrix} \begin{pmatrix} m \\ b \end{pmatrix}$$

$$\begin{pmatrix} m \\ b \end{pmatrix} = \begin{pmatrix} \sum_{i=1}^{N} x_i^2 & \sum_{i=1}^{N} x_i \\ \sum_{i=1}^{N} x_i & N \end{pmatrix}^{-1} \begin{pmatrix} \sum_{i=1}^{N} y_i x_i \\ \sum_{i=1}^{N} y_i \end{pmatrix}$$

7 Least squares (4)

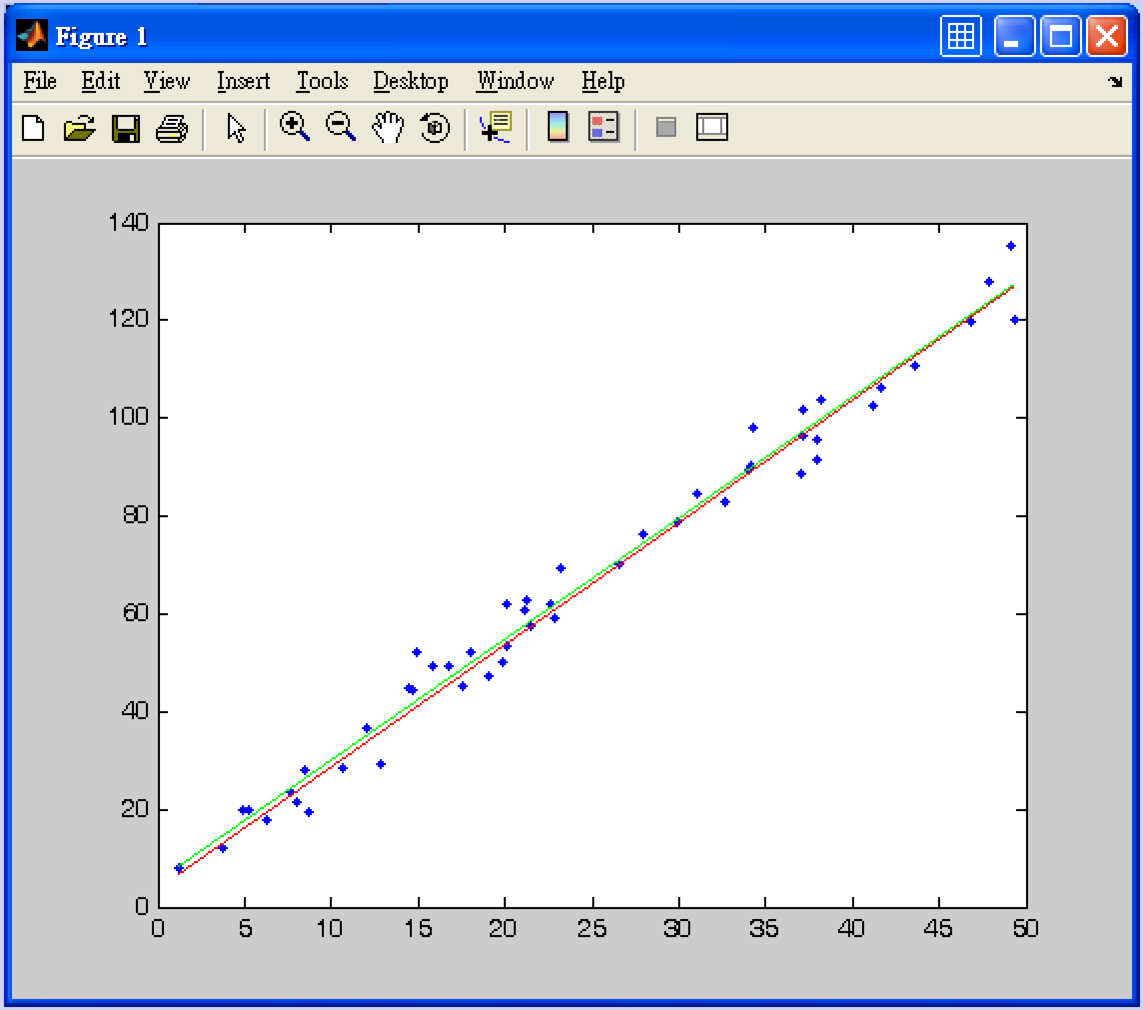
Matlab program:

```
m=2.5; b=3.7;
x=rand(1,50)*50;
y=m*x+b+randn(1,50)*5;
sx=[min(x) max(x)];
sy=m*sx+b;
figure(1); clf; plot(x,y,'.'); hold on; plot(sx,sy,'r');
A=[sum(x.^2)  sum(x);
   sum(x)     length(x)];
B=[sum(x.*y); sum(y)];
v=inv(A)*B ;
m1=v(1)
b1=v(2)

y1=m1*sx+b1;
plot(sx,y1,'g');
```

7 Least squares (5)

Result



Red: Ground truth, Green: Estimation.

Homework

- 計算f長度，$f = \sin t, \qquad |f| = \sqrt{\int_{-\pi}^{\pi} \sin^2 t \, dt}$