Vector representation of statistical problems.

Regression of Y on X $\Rightarrow$ $E(Y \mid X = x) = \int y\, f(y \mid x)\, dy$ which is a function of $\underline{x}$

If (X, Y) have joint distribution. $f(y \mid X = x) = \dfrac{f(x, y)}{f_X(x)}$

If X is non-stochastic then. $E(Y \mid x) = g(x, \beta) \rightarrow$ non-random.

data : $Y = g(x, \beta) + \text{error}$.

1 Simple linear regression:

Data : $y_i = \beta_0 + \beta_1 x_i + \epsilon_i$

data set $= \{(y_i, x_i) \quad i = 1, 2, \ldots n.\}$

$(\beta_0, \beta_1, \sigma^2)$ unknown parameters.

$\epsilon_i \underset{\sim}{\text{iid}} N(0, \sigma^2)$

x_i are non-stochastic.

y_i's are independent but NOT identically distributed.

$$\underset{\sim}{y} = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} \quad \underset{\sim}{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \quad \underset{\sim}{\beta} = \begin{pmatrix} \beta_0 \\ \beta_1 \end{pmatrix} \quad \underset{\sim}{\epsilon} = \begin{pmatrix} \epsilon_1 \\ \epsilon_2 \\ \vdots \\ \epsilon_n \end{pmatrix} \quad \underset{\sim}{1} = \begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix}.$$

$$\underset{n \times 1}{\underset{\sim}{y}} = \underset{n \times 2}{[\underset{\sim}{1} \ \underset{\sim}{x}]} \ \underset{2 \times 1}{\underset{\sim}{\beta}} + \underset{n \times 1}{\underset{\sim}{\epsilon}}$$

n - dimentional problem in $\mathbb{R}^n$

$\downarrow$

vectorspace.

2 weighted sum.

$\underset{\sim}{w} = (w_1, w_2 \cdots w_n)^T$

$\underset{\sim}{v} = (v_1, v_2 \cdots v_n)^T.$

$w^* = \dfrac{\underset{\sim}{w}}{\sum_{i=1}^{n} w_i}$ ~~w1~~,

$\underset{\sim}{w} \cdot \underset{\sim}{v} = \underset{\sim}{w}^T \underset{\sim}{v} = \sum_{i=1}^{n} w_i v_i$

$w_i^* \rightarrow$ credit.

$v_i \rightarrow$ Gr./Seml.

$(\underset{\sim}{w}^*) \cdot \underset{\sim}{v} = (\underset{\sim}{w}^*)^T \underset{\sim}{v} = \sum_{i=1}^{n} w_i^* v_i \longrightarrow$ <u>SGPA.</u>

dot product $\Leftarrow$ Inner product. $[\langle a, b \rangle]$

3 $\underset{\sim}{P} = \{ p_i \mid p_i \geq 0, \sum_{i=1}^{\infty} p_i = 1 \}.$

$\underset{\sim}{v} = \{ v_i = i \mid i \in \mathbb{N}. \}.$

$\langle \underset{\sim}{P}, \underset{\sim}{v} \rangle = \sum_{i=1}^{\infty} p_i v_i \rightarrow E(V).$

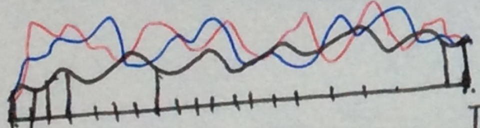
Convex combination. $\rightarrow$ Math.

Expectation R.V $\rightarrow$ Stat.

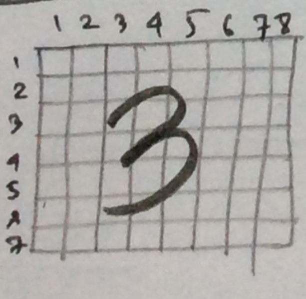
4 <u>High dimentional ~~d~~ data.</u>

Magnaticfield.

ECG ~~data~~.



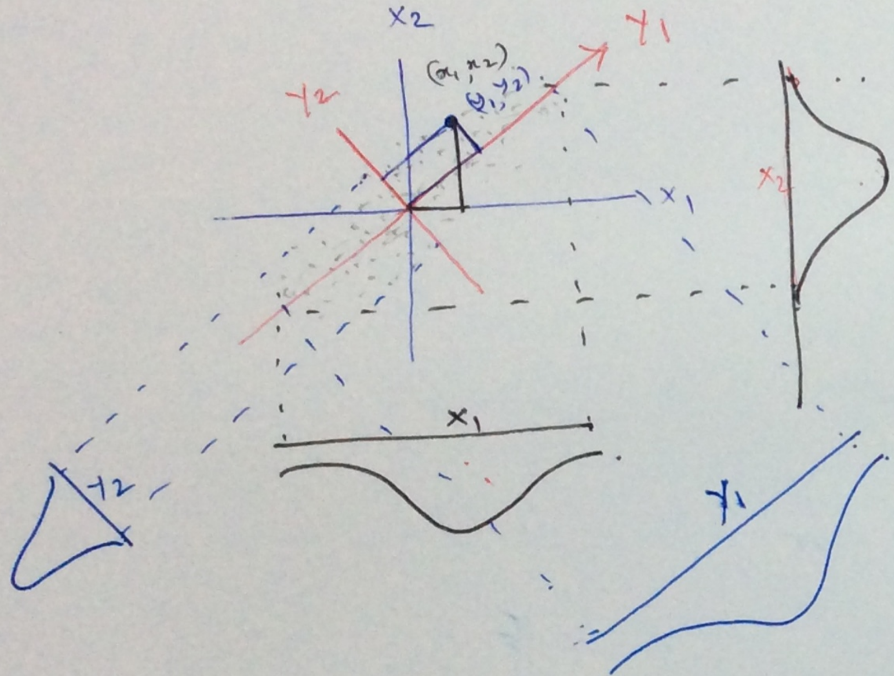
5. <u>Image analysis:</u>



$$\begin{bmatrix} I & c_1 \\ I & c_2 \\ I & c_3 \\ \vdots & \\ I & c_7. \end{bmatrix}$$

$7 \times 8 = \underline{56}.$

(6) Singular value decomposition (SVD) / Principal component analysis (PCA).



Coordinate change (Data driven)

The direction which has larger variation contributes more in model building and hence we incorporate the direction which have larger variance in the model.

(7) $\mathbb{P}_n = \{$ all polynomials of degree $\leq n \}$.

$$= \left\{ \sum_{i=0}^{n} a_i x^i \;\middle|\; a_i \in \mathbb{R} \right\}.$$

$$= \left\{ \underline{\tilde{a}}^T \tilde{X} \;\middle|\; \begin{array}{l} \tilde{a} = (a_0\, a_1 \cdots a_n)^T \\ \tilde{X} = (x^0, x^1 \cdots x^n)^T \end{array} \right\}.$$

Domain = $\mathbb{R}$.

$\tilde{a} \in \mathbb{R}^{n+1}$

$\Downarrow$

characterizes the polynomial.

Examples of vector space:

(i) $\mathbb{R}$.

(ii) $\mathbb{C}$.

(iii) $\mathbb{R}^k$

(iv) $\mathbb{P}_n = \left\{ \sum_{i=0}^{n} a_i x^i \quad a_i \in \mathbb{R}, \; x \in \mathbb{R} \right\}$.

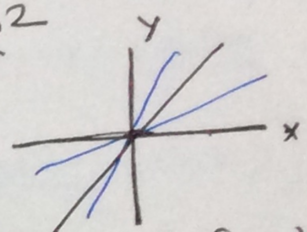
(v) All continuous function on $[a, b]$.

(*) Intersection of more than one subspaces is a sub-space.

(*) Union of more than one subspaces is NOT a sub-space.

Examples of sub-spaces.

(i) $V = \mathbb{R}^k = \mathbb{R}^2$



→ $\{(0,0)\}$

→ x axis

→ y axis.

→ any line passing through (0, 0).

→ $\mathbb{R}^2$

(ii) $V = \mathbb{R}^3$

→ $\{(0, 0, 0)\}$

→ x axis, y axis, z axis.

→ any line passing through (0, 0, 0)

→ any plane passing through (0, 0, 0).

→ $\mathbb{R}^3$.

(iii) $V = \mathbb{P}_n$

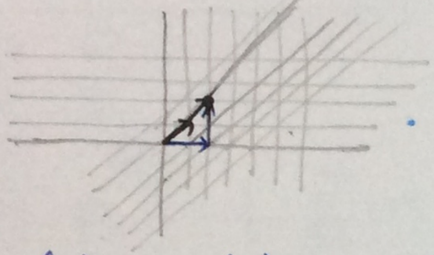
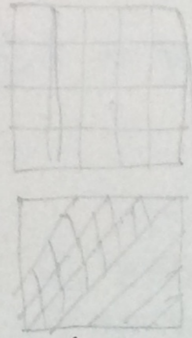
Any $\mathbb{P}_m$ where $m \le n$.

$(a_0 \, a_1 \cdots a_n) \in \mathbb{R}^{n+1}$.

$(a_0 \, a_1 \cdots a_{n-1} \, 0) \to \mathbb{P}_{n-1}$

$(a_0 \, a_1 \cdots a_{n-2} \, 0 \, 0) \to \mathbb{P}_{n-2}$

Example of span.

(1)  

$$\begin{pmatrix}2\\2\end{pmatrix} = 2\begin{pmatrix}1\\0\end{pmatrix} + 2\begin{pmatrix}0\\1\end{pmatrix} \to \begin{pmatrix}2\\2\end{pmatrix}$$

$$\begin{pmatrix}2\\2\end{pmatrix} = 0\begin{pmatrix}0\\1\end{pmatrix} + 2\begin{pmatrix}1\\1\end{pmatrix} \to \begin{pmatrix}0\\2\end{pmatrix}$$

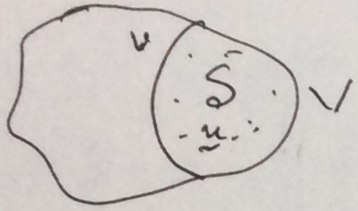
$$sp\left\{\begin{pmatrix}1\\1\end{pmatrix}, \begin{pmatrix}0\\1\end{pmatrix}\right\} = sp\left\{\begin{pmatrix}1\\0\end{pmatrix}, \begin{pmatrix}0\\1\end{pmatrix}\right\} = \mathbb{R}$$

(2) $$sp\left\{\begin{pmatrix}1\\1\\0\end{pmatrix}, \begin{pmatrix}0\\1\\0\end{pmatrix}, \begin{pmatrix}0\\0\\0\end{pmatrix}\right\} = \mathbb{R}\times\mathbb{R}\times\{0\} \neq \mathbb{R}^2 \subset \mathbb{R}^3$$

$\mathbb{R}\times\mathbb{R}\times\{0\}$ is ishomorphic to $\mathbb{R}^2$

(*) **Remark:** If a collection of n-component vectors consists of more than n vectors then they are linearly dependant.

Example: $\left\{\begin{pmatrix}1\\0\\0\end{pmatrix}, \begin{pmatrix}0\\1\\0\end{pmatrix}\right\}$ is a basis of $S = \mathbb{R}\times\mathbb{R}\times\{0\}$ sub-space of $\mathbb{R}^3 = V$.



$$\{ v \in V \mid v^T x = 0 \quad \forall x \in S \} = S^{\perp}.$$

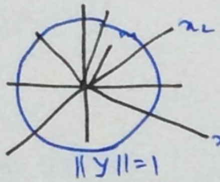
$$S \cap S^{\perp} = 0.$$

$V = \mathbb{R}^2$ $\quad S =$ x-axis. $\quad S^{\perp} =$ y axis.

$\hat{y}$ and $y - \hat{y}$

$\hat{y} \downarrow$ Prediction

$y - \hat{y} \downarrow$ error.

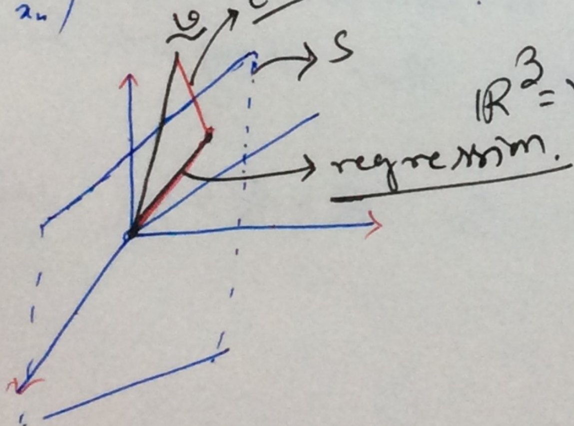


~~$x_1 x_2 \dots x_n$~~. $\quad x = (x_1, x_2, \dots, x_n)^T \longrightarrow y = \dfrac{x}{\|x\|} = \dfrac{x}{\sqrt{x^T x}} = \dfrac{x}{(\sum x_i^2)^{1/2}}.$

if $x \neq 0$.

$$x = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$$

$\vec{x}$ $\quad \bar{x}$ mean



$\mathbb{R}^3 = V.$

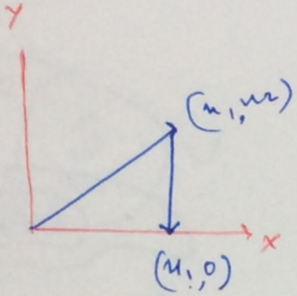
$y \in \mathbb{R}^n.$

$y \to \hat{y} \in S = sp\{1, x\}.$

$y \to y - \hat{y} \in S^{\perp}.$

$\hat{y} = M y.$

??



$$\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}\begin{pmatrix} u_1 \\ u_2 \end{pmatrix} = \begin{pmatrix} u_1 \\ 0 \end{pmatrix}$$

$$I_2 = P_S + P_{S^\perp}$$

$V = \mathbb{R}^2$

$S = x$ axis.

$P_S = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$.

$P_{S^\perp} = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$

$P_S \underset{\sim}{u} = \begin{pmatrix} u_1 \\ 0 \end{pmatrix}$

$P_S \begin{pmatrix} u_1 \\ 0 \end{pmatrix} = \begin{pmatrix} u_1 \\ 0 \end{pmatrix}$

$P_{S^\perp} \begin{pmatrix} u_1 \\ u_2 \end{pmatrix} = \begin{pmatrix} 0 \\ u_2 \end{pmatrix}$

$\forall \underset{\sim}{v} \in V.$

$S \subseteq V$

$P_S \underset{\sim}{v} \in S. \ \forall \underset{\sim}{v} \in V.$ ✓

$\Rightarrow P_S (P_S \underset{\sim}{v}) = P_S \underset{\sim}{v}. \ \forall \underset{\sim}{v} \in V.$ ✓

$\Rightarrow \underline{P_S^2 \underset{\sim}{v} = P_S \underset{\sim}{v} \quad \forall \underset{\sim}{v} \in V.}$ ✓

$\Downarrow$.

$\boxed{P_S^2 = P_S.}$ $\Rightarrow$ <u>Idempotent matrix.</u> ✓✓

<u>given</u>

$P^2 = P$

$P \underset{\sim}{v} = \lambda \underset{\sim}{v}. \qquad \underset{\sim}{v} \neq \underset{\sim}{0}.$

$\Rightarrow P P \underset{\sim}{v} = \lambda P \underset{\sim}{v}$

$P^2 \underset{\sim}{v} = \lambda^2 v$

$\Rightarrow P \underset{\sim}{v} = \lambda^2 \underset{\sim}{v}$

$\Rightarrow \lambda \underset{\sim}{v} = \lambda^2 \underset{\sim}{v}.$

$\Rightarrow (\lambda^2 - \lambda) \underset{\sim}{v} = \underset{\sim}{0}$.

$\underset{\sim}{v}^T \underset{\sim}{v} \neq 0$ as $\underset{\sim}{v} \neq 0$

$(\lambda^2 - \lambda) \underset{\sim}{v}^T \underset{\sim}{v} = \underset{\sim}{v}^T \underset{\sim}{0} = 0$

$\Rightarrow \lambda^2 - \lambda = 0$

$\Rightarrow \lambda = 0, 1$

Idempotent matrix has eigen values <u>0, 1 only</u>.

$$\left[\begin{array}{cc:c} 1 & 2 & 3 \\ 1 & 3 & 5 \end{array}\right]$$
$A \quad B$

$$\rho(AB) = \begin{cases} \rho(A) & \text{if } B = I. \\ \underset{\sim}{0}. & \text{if } B = 0 \end{cases}$$

④ $\rho(AA^T) = \rho(A).$

① $\rho(AA^T) \subseteq \rho(A) \longrightarrow$ By the property(2).

② To show. $\rho(AA^T) \supseteq \rho(A).$

Let $\underset{\sim}{l} \in [\rho(AA^T)]^{\perp}.$ $\quad \underset{\sim}{l} \neq \underset{\sim}{0}.$

$\Leftrightarrow \underset{\sim}{l}^T [AA^T] = \underset{\sim}{0}^T$ — ⊛

$\Leftrightarrow \underset{\sim}{l}^T AA^T \underset{\sim}{l} = \underset{\sim}{0}^T \underset{\sim}{l}$

$\Leftrightarrow (A^T \underset{\sim}{l})^T (A^T \underset{\sim}{l}) = 0$

$\Leftrightarrow \underset{\sim}{u}^T \underset{\sim}{u} = 0$

$\Leftrightarrow \underset{\sim}{u} = \underset{\sim}{0}$

$\Leftrightarrow A^T l = \underset{\sim}{0}$

$\Leftrightarrow l^T A = \underset{\sim}{0}^T$ — ⊛

$\Leftrightarrow l \in [\rho(A)]^{\perp}.$

$A^T \underset{\sim}{l} = \underset{\sim}{u}$

$u^T u = \sum_{i=1}^{n} u_i^2 \quad u_i \in \mathbb{R}.$

$$[\rho(AA^T)]^{\perp} \subseteq [\rho(A)]^{\perp}.$$

$$\rho(A) \subseteq \rho(AA^T).$$

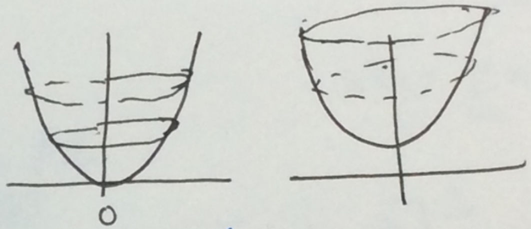
$$\underset{\sim}{x}^T A \underset{\sim}{x} = \sum_{i=1}^{n} \sum_{j=1}^{n} x_i x_j a_{ij} \rightarrow \text{Quadratic form.}$$

$$A = I \Rightarrow x^T I \underset{\sim}{x} = \sum_{i=1}^{n} x_i^2 \rightarrow \text{Euclidian distance.}$$

(I) If A is p.d. then all eigen values of A are positive.

(II) $|A| = \prod_{i=1}^{n} \lambda_i > 0.$

(III) $Tr(A) = \sum_{i=1}^{n} \lambda_i > 0.$



$x_i \sim N(0, \sigma_i^2)$ and they are independent.

$$\Rightarrow \frac{x_i}{\sigma_i} \sim N(0,1) \Rightarrow \sum_{i=1}^{n} \left(\frac{x_i}{\sigma_i}\right)^2 \sim \chi^2_n$$

$$\Rightarrow \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}^T \begin{bmatrix} 1/\sigma_1^2 & & 0 \\ & 1/\sigma_2^2 & \\ 0 & & 1/\sigma_n^2 \end{bmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \sim \chi^2_n.$$

$$\boxed{Z = \frac{X - E(x)}{sd(x)}}$$

$$\Rightarrow \underset{\sim}{X}^T D \underset{\sim}{X} \sim \chi^2_n.$$

$$D^2 = (\underset{\sim}{x} - \underset{\sim}{\mu})^T \Sigma^{-1} (\underset{\sim}{x} - \underset{\sim}{\mu})$$

→ Mohalanobish D^2 distance.