$X \in \mathbb{R}^{N \times d}$

$$X = \begin{bmatrix} x_{11} & \cdots & x_{1d} \\ \vdots & & \\ x_{N1} & \cdots & x_{Nd} \end{bmatrix}_{N \times d}$$

x, y : random variable

$$\operatorname{cov}(x,y) = E[(x-\mu_x)(y-\mu_y)] \leftarrow \text{True covariance}$$

$$= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (\lambda_1 - \mu_x)(\lambda_2 - \mu_y) f_{x,y}(\lambda_1, \lambda_2) \, d\lambda_1 \, d\lambda_2$$

$f_{x,y}(\lambda_1, \lambda_2)$: Joint density

Covariance Approximation knowing Data Matrix X

Step 1: $\tilde{X} = X - \bar{X}$

Step 2: $H = \frac{\tilde{X}^T \tilde{X}}{N-1}$ = Covariance Matrix

$H \in \mathbb{R}^{d \times d} \longrightarrow$ square matrix / symmetric $H^T = H$

variance

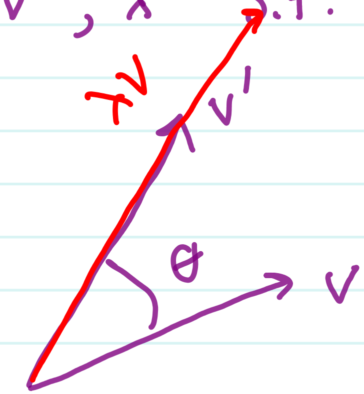
$$H = \begin{bmatrix} \Sigma_{x_1 x_1} & \Sigma_{x_1 x_2} & \cdots & \Sigma_{x_1 x_d} \\ \Sigma_{x_2 x_1} & \Sigma_{x_2 x_2} & & \vdots \\ \vdots & & \ddots & \\ \Sigma_{x_d x_1} & & & \Sigma_{x_d x_d} \end{bmatrix}$$

$\Sigma_{x_1 x_2} = \Sigma_{x_2 x_1}$

Finding eigen values / eigen vectors.

$A \in \mathbb{R}^{n \times n} \rightarrow$ square

e-values : Find v, λ s.t. $\boxed{Av = \lambda v}$, $v \neq 0$



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

$v' = A \cdot v$

$Av = \lambda v \rightarrow Av - \lambda v = 0 \rightarrow v(A - \lambda I) = 0$

$v \neq 0$ is in the null space of $(A - \lambda I)$ space

Ex :

$$A = \begin{bmatrix} 7 & 4 \\ 4 & 1 \end{bmatrix}$$

Find e-value and e-vector of A. Solve

$$Av = \lambda v$$

characteristic equation

$$v(A - \lambda I) = 0 \longrightarrow |\lambda I - A| = 0$$

$$\left| \lambda \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} - \begin{bmatrix} 7 & 4 \\ 4 & 1 \end{bmatrix} \right| = 0 \qquad \begin{vmatrix} \lambda - 7 & -4 \\ -4 & \lambda - 1 \end{vmatrix} = 0$$

$$(\lambda - 7)(\lambda - 1) - 16 = 0 \longrightarrow \lambda^2 - 8\lambda + 7 - 16 = 0$$

$$\lambda^2 - 8\lambda - 9 = 0$$

$$ax^2 + bx + c = 0$$

$$x = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}$$

$$\Rightarrow (\lambda+1)(\lambda-9)=0 \qquad \begin{cases} \lambda_1 = -1 \\ \lambda_2 = 9 \end{cases}$$

— Finding e-vector for $\lambda_1 = -1$

$$AV_1 = \lambda_1 V_1 \rightarrow \begin{bmatrix} 7 & 4 \\ 4 & 1 \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} -x \\ -y \end{bmatrix}$$

$$\begin{cases} 7x + 4y = -x \\ 4x + y = -y \end{cases} \rightarrow 8x = -4y \rightarrow -2x = y$$

$$V_1 = \begin{bmatrix} 1 \\ -2 \end{bmatrix}$$

$$\lambda_1 = -1 \quad , \quad V_1 = \begin{bmatrix} 1 \\ -2 \end{bmatrix}$$

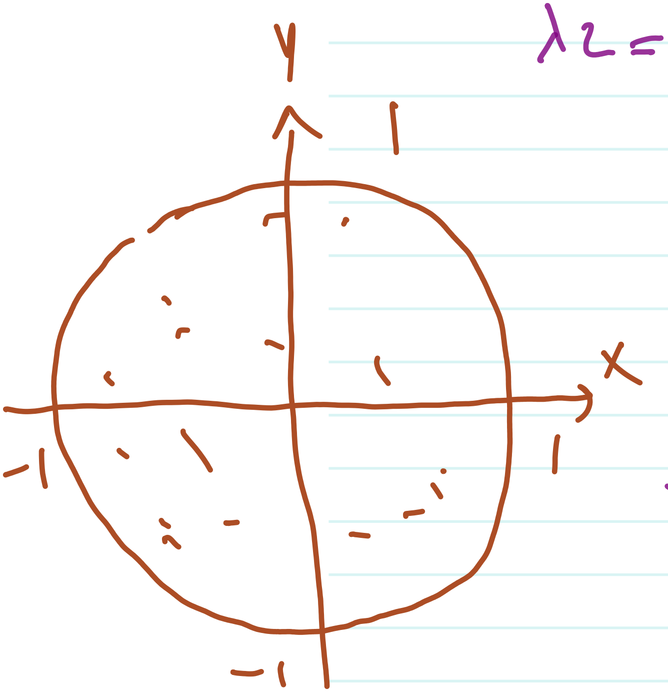
— Finding e-vector for $\lambda_2 = 9$

$$A V_2 = \lambda_2 V_2 \longrightarrow \begin{bmatrix} 7 & 4 \\ 4 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 9x \\ 9y \end{bmatrix}$$

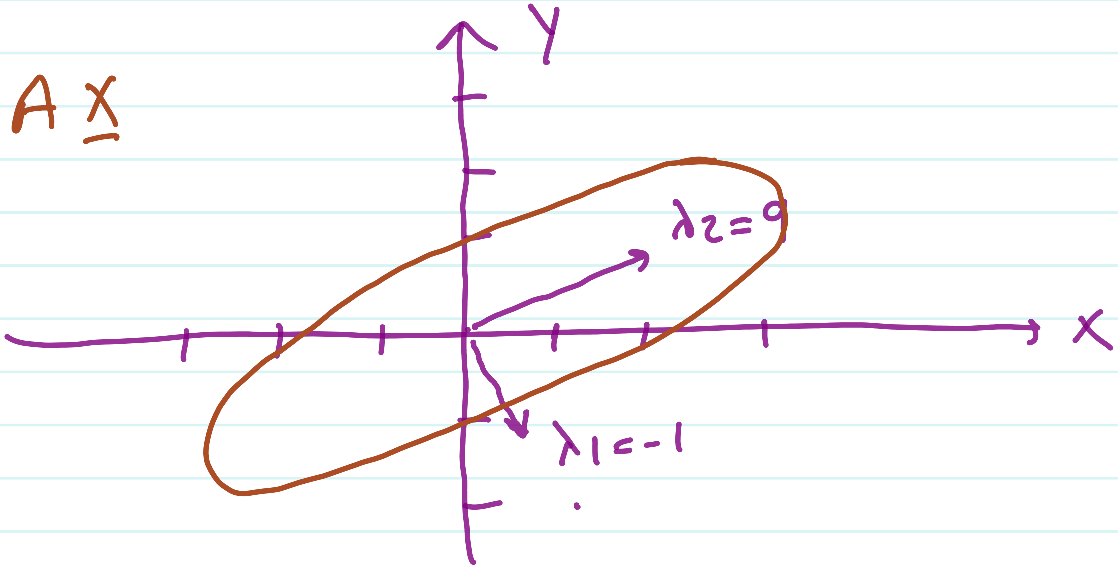
$$\begin{cases} 7x + 4y = 9x \\ 4x + y = 9y \end{cases} \longrightarrow 2x = 4y \longrightarrow x = 2y$$

$$V_2 = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$$

$$\lambda_2 = 9 \ , \quad V_2 = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$$



$A\underline{X}$



Suppose $\underline{V} = \begin{bmatrix} \underline{v}_1 & \underline{v}_2 & \cdots & \underline{v}_d \end{bmatrix}$

$\underline{v}_i$ is eigen vector of X

$$\underline{V} = \begin{bmatrix} 1 & 2 \\ -2 & 1 \end{bmatrix}$$

$$V^{-1} A V = \begin{bmatrix} 1 & 2 \\ -2 & 1 \end{bmatrix}^{-1} \begin{bmatrix} 7 & 4 \\ 4 & 1 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ -2 & 1 \end{bmatrix}$$

$$= \begin{bmatrix} 1 & 2 \\ -2 & 1 \end{bmatrix}^{-1} \begin{bmatrix} -1 & 18 \\ 2 & 9 \end{bmatrix}$$

$$= \frac{\begin{bmatrix} 1 & -2 \\ 2 & 1 \end{bmatrix}}{1+4} \begin{bmatrix} -1 & 18 \\ 2 & 9 \end{bmatrix}$$

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$$

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

$$= \frac{1}{5}\begin{bmatrix} -5 & 0 \\ 0 & 45 \end{bmatrix} = \begin{bmatrix} -1 & 0 \\ 0 & 9 \end{bmatrix}$$

$$V^{-1} A V = \Lambda = \begin{bmatrix} \lambda_1 & & & 0 \\ & \lambda_2 & & \\ & & \ddots & \\ 0 & & & \lambda_d \end{bmatrix}$$

reciprocal basis

diagonalization

$$A = \begin{bmatrix} 7 & 4 \\ 4 & 1 \end{bmatrix}$$

$$f(x) = A^n x$$

$$= A\,(A^{n-1})\,x$$

$$= A \cdot A\,(A^{n-2})\,x$$

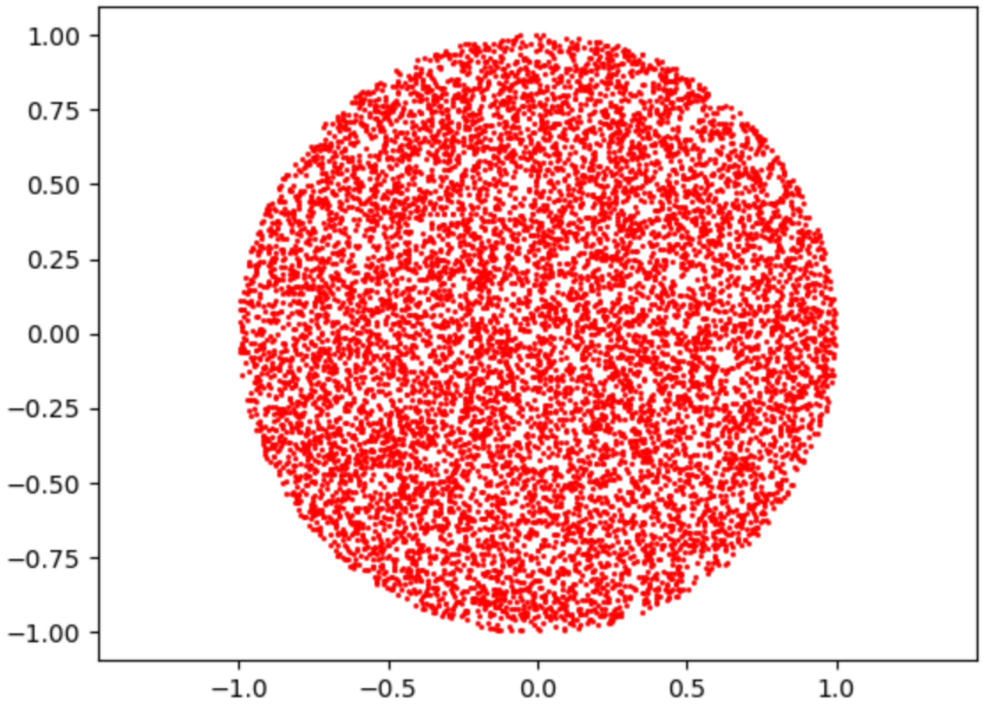
$$Ax = \begin{bmatrix} 7 & 4 \\ 4 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}$$

$$\Lambda x = \begin{bmatrix} -1 & 0 \\ 0 & 9 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} -x \\ 9y \end{bmatrix}$$

$$\Lambda^2 x = \Lambda(\Lambda x) = \Lambda\left(\begin{bmatrix} -x \\ 9y \end{bmatrix} \right) = \begin{bmatrix} x \\ 9^2 y \end{bmatrix}$$

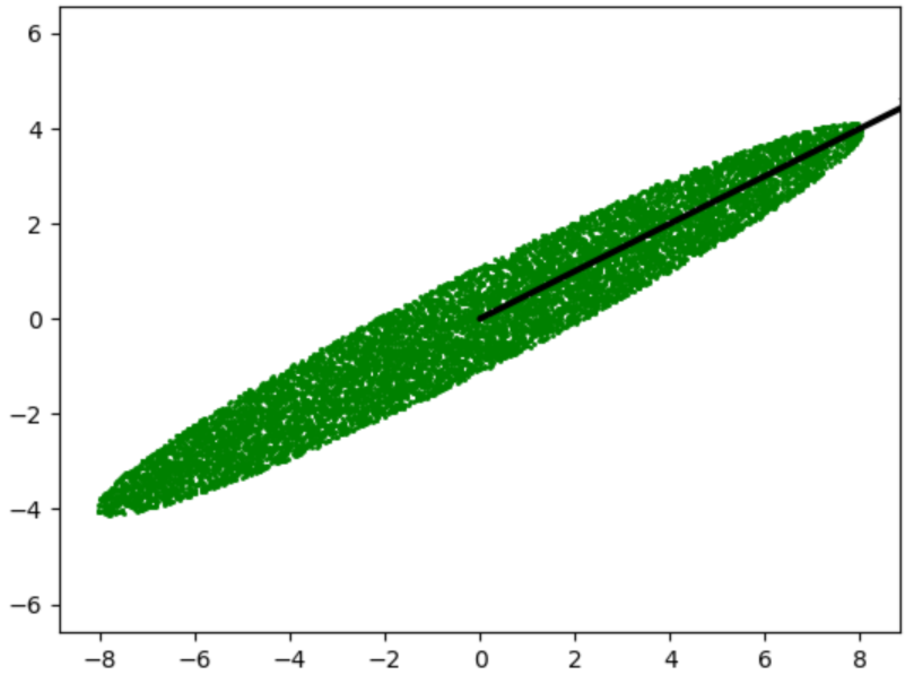
$$\Lambda = \begin{bmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_d \end{bmatrix} \rightarrow \Lambda^n = \begin{bmatrix} \lambda_1^n & & & \\ & \lambda_2^n & & \\ & & \ddots & \\ & & & \lambda_d^2 \end{bmatrix}$$

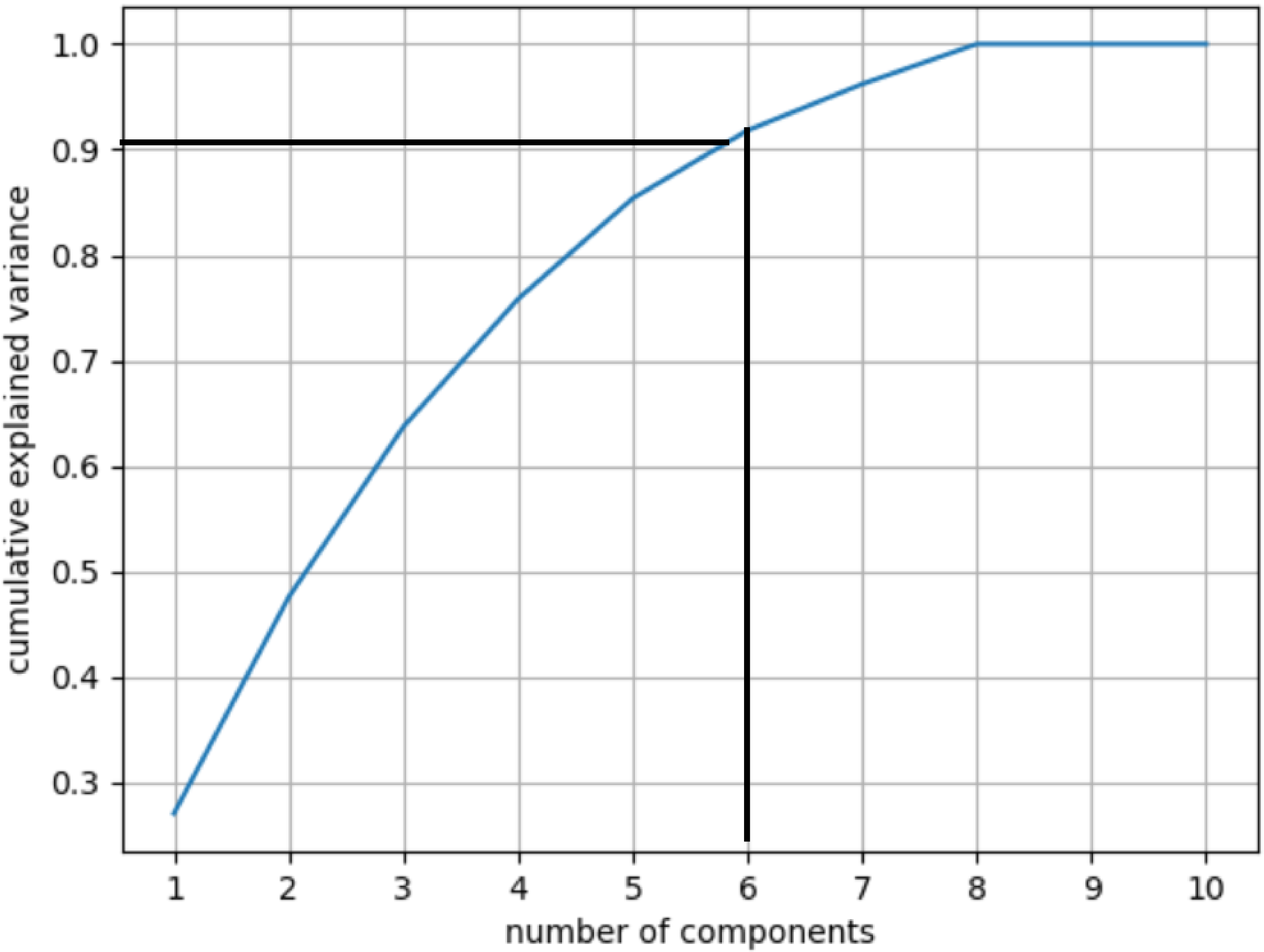
$$\Lambda^{-1} = \begin{bmatrix} \frac{1}{\lambda_1} & \ddots & 0 \\ 0 & & \frac{1}{\lambda_d} \end{bmatrix}$$



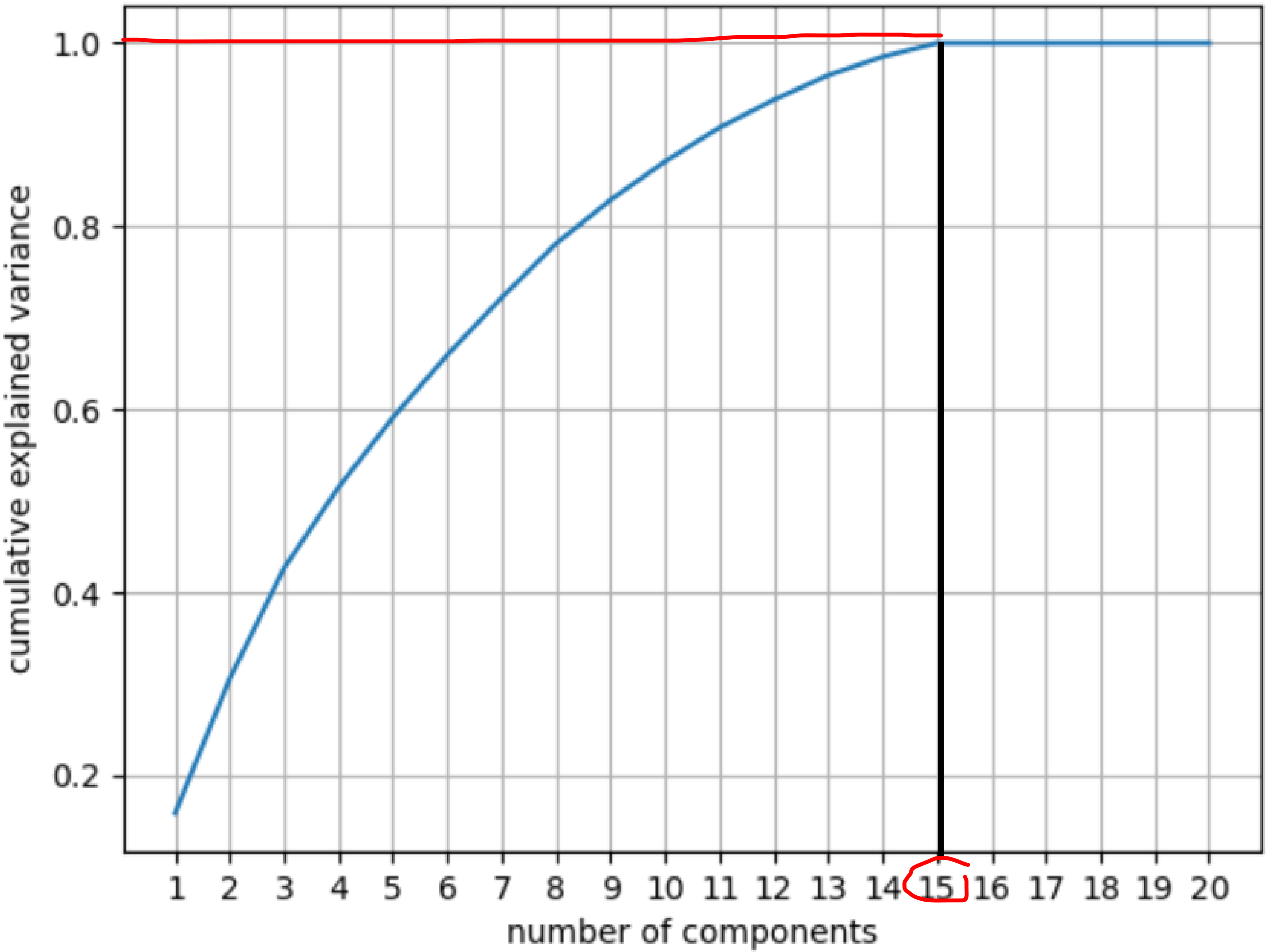
$$A = \begin{bmatrix} 7 & 4 \\ 4 & 1 \end{bmatrix}$$

$$A \underline{x}$$





cumulative explained variance
1.0
0.9
0.8
0.7
0.6
0.5
0.4
0.3
1
2
3
4
5
6
7
8
9
10
number of components



1.0
0.8
0.6
0.4
0.2
cumulative explained variance
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20
number of components

$\boxed{\text{dummy Trap}}$

	gender
1	female
2	male
3	male
4	female

$\rightarrow$

	x_1 gender_f	x_2 gender_m
1	1	0
2	0	1
3	0	1
4	1	0

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 \qquad x_1 = 1 - x_2$$

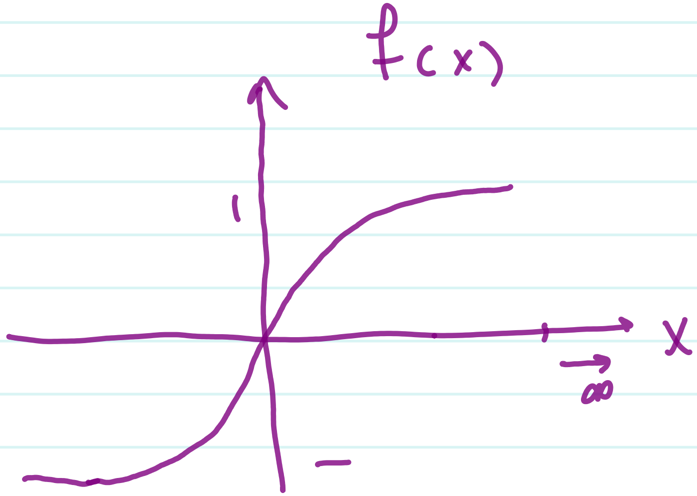
$$y = \beta_0 + \beta_1 (1 - x_2) + \beta_2 x_2$$

$\underline{x_1}$ $\downarrow$ $\underline{x_2}$

$$\beta^* = (x^T x)^{-1} x^T y \rightarrow x = \begin{bmatrix} 1 & x_{11} & x_{12} \\ \vdots & \vdots & \vdots \\ 1 & x_{N1} & x_{2Ni} \end{bmatrix}$$

M — 1

Sun = 1 → (M, T, W, R, S)

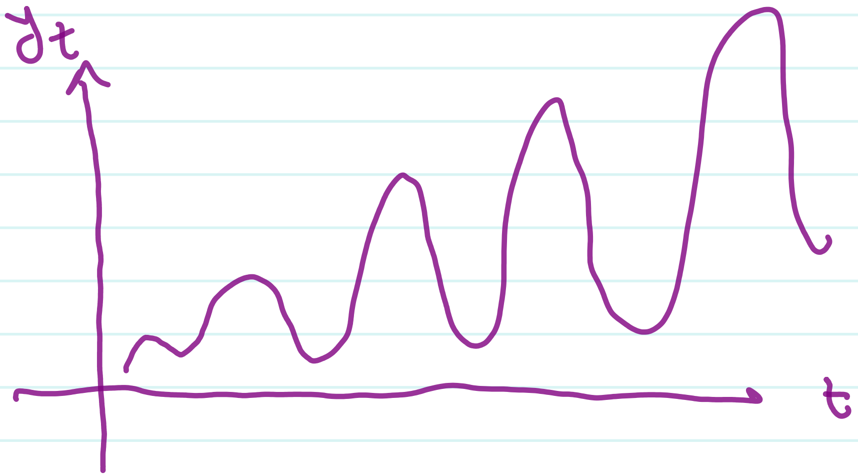


$f(x) = \tanh(x)$

$$\lim_{x \to \infty} f(x) = 1$$

$$\lim_{x \to -\infty} f(x) = -1$$

$$\{y_i\}_{i=1}^{N} = \{y_1, y_2, \dots y_n\}$$



non-stationary data set

find Model

$$\{z_i = \log y_i\}_{i=1}^{N}$$

$z_1 = \log y_1$

$z_2 = \log y_2$

$\vdots$

$$\{z_i = \sqrt{y_i}\}_{i=1}^{N}$$

$\hat{z}_{t+h|t} \rightarrow \hat{y}_{t+h|t}$

reverse transformation $\leftarrow$ $\hat{y}_{t+h|t} = e^{\hat{z}_{t+h|t}}$

differencing

y	$z = \Delta y = y_t - y_{t-1}$	$\Delta^2 y = \Delta z = m$
y_1	—	—
y_2	$z_1 = y_2 - y_1$	—
y_3	$z_2 = y_3 - y_2$	$m_1 = z_2 - z_1$
y_4	$z_3 = y_4 - y_3$	$m_2 = z_3 - z_2$
y_5	$z_4 = y_5 - y_4$	$m_3 = z_4 - z_3$

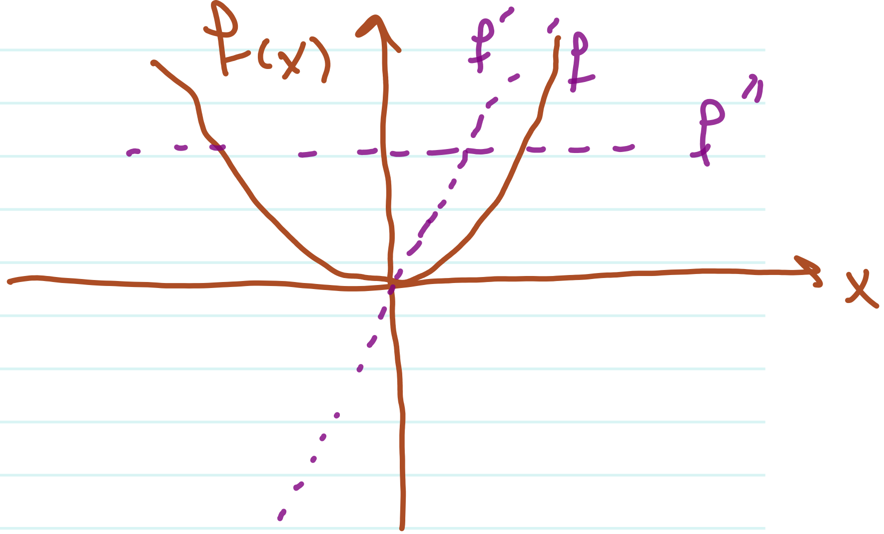
$\hat{y} \leftarrow \hat{z}$

differencing helps to make non-stationary data into stationary if trend

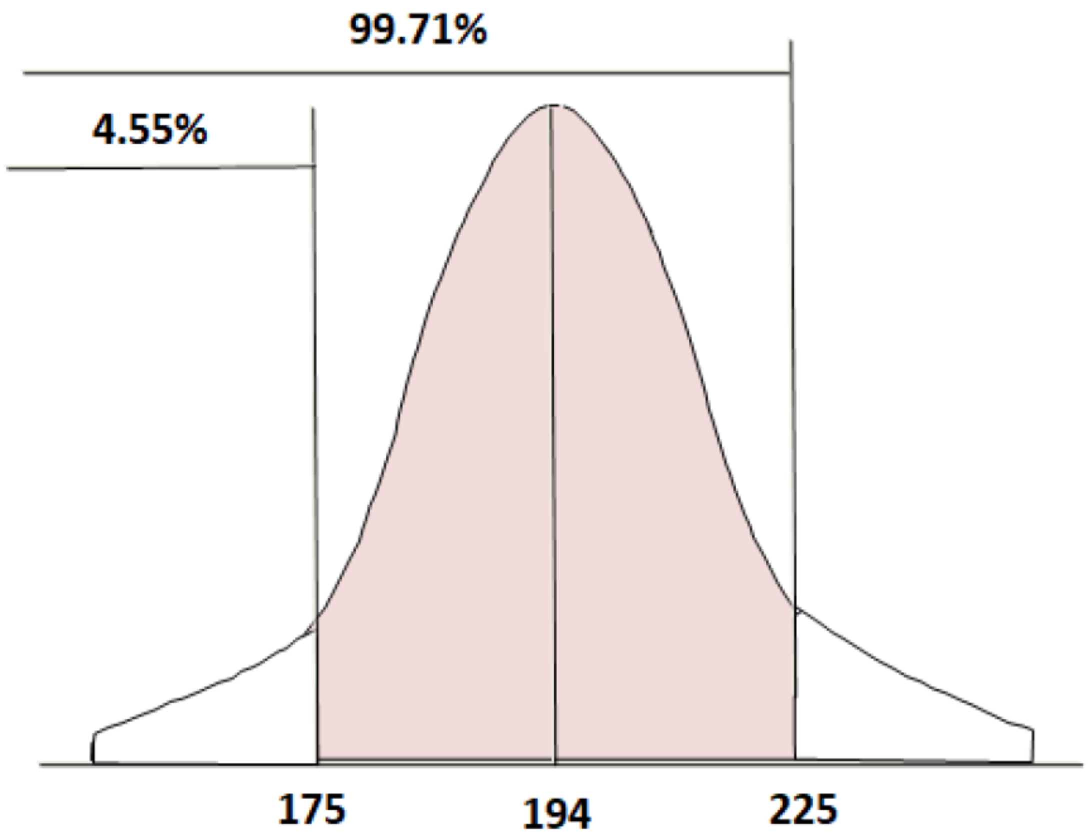
$f(x) = x^2$

$f'(x) = 2x$

$f''(x) = 2$



- The weights of 500 American men were taken and the sample mean was found to be 194 pounds with a standard deviation of 11.2 pounds. What percentages of men have weights between 175 and 225 pounds?



$\mu = 194$

$\sigma = 11.2$

$$z_1 = \frac{175 - 194}{11.2}$$

$$z_2 = \frac{225 - 194}{11.2}$$

$=$