10601

Hsueh Lin Huang

hsuehlih

Vectors and Matrices

$$X = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} \quad y = \begin{pmatrix} 1 \\ 2 \end{pmatrix} \quad z = \begin{pmatrix} 3 \\ 4 \end{pmatrix}$$

1. $y^T z = \begin{pmatrix} 1 & 2 \end{pmatrix} \begin{pmatrix} 3 \\ 4 \end{pmatrix} = 11$

2. $Xy = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} \begin{pmatrix} 1 \\ 2 \end{pmatrix} = \begin{pmatrix} 5 \\ 11 \end{pmatrix}$

3. $\det X = 1 \cdot 4 - 3 \cdot 2 = -2$, yes, it is invertible.

$$\frac{\begin{pmatrix} 4 & -2 \\ -3 & 1 \end{pmatrix}}{-2} = \begin{pmatrix} -2 & 1 \\ \frac{3}{2} & -\frac{1}{2} \end{pmatrix} \#$$

4. rank of $X = 2$

Calculus

1. $y = x^4 + 2x^2 - 1$

$y' = 4x^3 + 4x$ #

2. $y = \log\left(\frac{x^7}{10x}\right) + \sin(z) x^{z-8} = \frac{6}{x \cdot \ln(10)} + (z-8)\sin(z) x^{z-9}$ #

Probability and Statistics

1. $\frac{0+1+1+0+0}{5} = 0.4$ #

2. $s^2 = \frac{(0-0.4)^2 + (1-0.4)^2 + (1-0.4)^2 + (0-0.4)^2 + (0-0.4)^2}{4} = 0.3$

3. $0.4 \times 0.6 \times 0.6 \times 0.4 \times 0.4 = 0.02304$ #

4. $P = (1-x)^3 x^2$

$P' = 3(1-x)^2(-1)x^2 + (1-x)^3 \cdot 2x = (1-x)^2 x (2-5x) = 0, \; x = 0, 1, \frac{2}{5}$

if $x = 0$ or $1 \Rightarrow P = 0$, $x = \frac{2}{5} \Rightarrow P = 0.03456$ (max) #

5.

(a) $P(A=0, B=0) = 0.5$

(b) $P(A=1) = 0.4$

(c) $P(A=0 | B=1) = 0.5$

(d) $P(A=0 \vee B=0) = 0.9$

Big-O Notation

1. $f(n) = 2^n$, $g(n) = e^n$

$f(n) = O(g(n))$ Assume $n > 1$. $e^n > 2^n$

2. $f(n) = n^2$, $g(n) = n^4 + 2n + 3$

$f(n) = O(g(n))$ n^4 dominates.

3. $f(n) = n$, $g(n) = \log_{10} n$

$g(n) = O(f(n))$ $n > \log n$, n dominates.

Algorithms

Use binary search. By comparing the target value to the value of the middle element of the sorted array. If the target value is equal to the middle element's value, then the position is returned and the search is finished. If the target value is greater / lesser than the middle element's value, then the search continues on the upper / lower half of the array until the target value is found or the entired array has been searched. (not found)

pseudo code.

```
BinaraySearch (array, target, l, r)
  if (l > r) return NULL;
  mid = l + (r-l)/2
  if (array[mid] == target)
      return mid;
  else if (array[mid] > target)
      return BS (array, target, l, mid-1)
  else
      return BS(array, target, mid+1, r)
```

Running time of Binary search

Master Theorem = Case 2

$n^{\log_b a} = n^0 = 1$

$f(n) = \Theta(1) = \Theta(n^{\log_b a} \lg^k n)$

$T(n) = \Theta(n^{\log_b a} \lg^{k+1} n) = \Theta(\lg n)$ *

Probability and Random Variables

Probability

1. true
2. false
3. false
4. false
5. false

Discrete and Continuous Distributions

1. 1d Gaussian distribution

(pdf) $N(x|\mu, \sigma) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{\frac{1}{2\sigma^2}(x-\mu)^2}$

2. Bernoulli (p) $0<p<1$

(pmf) $\begin{cases} q = (1-p) & \text{for } k=0 \\ p & \text{for } k=1 \end{cases}$

3. Uniform Unif (a, b), $a<b$

(pdf) $\begin{cases} \frac{1}{b-a} & \text{for } x \in [a,b] \\ 0 & \text{otherwise} \end{cases}$

4. Exponential Distribution

Exp(λ), $\lambda > 0$

(pdf) $\lambda e^{-\lambda x}$

5. Poisson Distribution

Poisson (λ), $\lambda>0$

(pmf) $\frac{\lambda^k e^{-\lambda}}{k!}$

Mean and Variance

1.

$$p(x;\lambda) = \frac{(\frac{\lambda}{3})^x e^{-\lambda/3}}{x!}, \lambda > 0$$

(a) $\frac{\lambda}{3}$

(b) $\frac{\lambda}{3}$

2.

(a) $E[2X] = 2$

(b) $Var(2X) = 4Var(X) = 4$

$Var(X+1) = Var(X) = 1$

Mutual and Conditional Independence

1. $Var(X+Y) = Var(X) + Var(Y)$

Proof: $Var(X+Y) = Var X + 2Cov(X,Y) + Var Y)$

$$= E[(X-E(X))^2 + (Y-E(Y))^2 + 2(X-E(X))(Y-E(Y))]$$

$$= \underbrace{E(X-E(X))^2}_{Var(X)} + \underbrace{E(Y-E(Y))^2}_{Var(Y)} + 2\underbrace{E((X-E(X))(Y-E(Y)))}_{cov(X,Y)}$$

$$= \underline{E(XY)} - E(X)E(Y)$$

$$= E(X)E(Y) - E(X)E(Y) = 0$$

2.

(a) Yes, X and Y are independent.

(b) No, because $X+Y = Z$ (even)

If X is an even number, then Y must be even, too,

If X is an odd number, then Y must be odd, too.

of Large Numbers and the Central Limit Theorem.

1.

$$\sqrt{n}\left(\left(\frac{1}{n}\sum_{i=1}^{n} X_i\right) - \mu\right) \xrightarrow{d} N(0,\sigma^2) \Rightarrow \sum_{i=1}^{n} X_i = N(\mu, \sqrt{n}\sigma^2)$$

mean $= 60000 \times \frac{1}{6} = 10000$ ← dominates, since variance is much smaller than mean.

variance $= \sqrt{60000} \times \frac{5}{36} = 34.02$

2. mean $= n \times \mu$

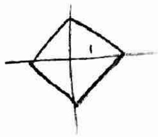
Variance $= \sqrt{n}\sigma^2$

If n is large enough, mean dominates the distribution.

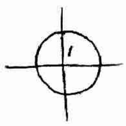
Linear Algebra

Vector Norms

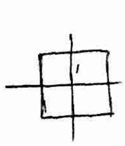
1. $\|x\|_1 \leq 1$



2. $\|x\|_2 \leq 1$



3. $\|x\|_\infty$



Geometry

1.

$$d = \frac{|Ax_1 + By_1 + Cz_1 + D|}{\sqrt{A^2+B^2+C^2}}, \quad Ax+By+Cz+D=0$$

origin $= (0,0,0)$

$$d = \frac{|A\cdot 0 + B\cdot 0 + C\cdot 0 + D|}{\sqrt{A^2+B^2+C^2}},$$

therefore, the distance between hyperplane $w^Tx + b = 0$ is $\frac{|b|}{\|w\|_2}$, true #

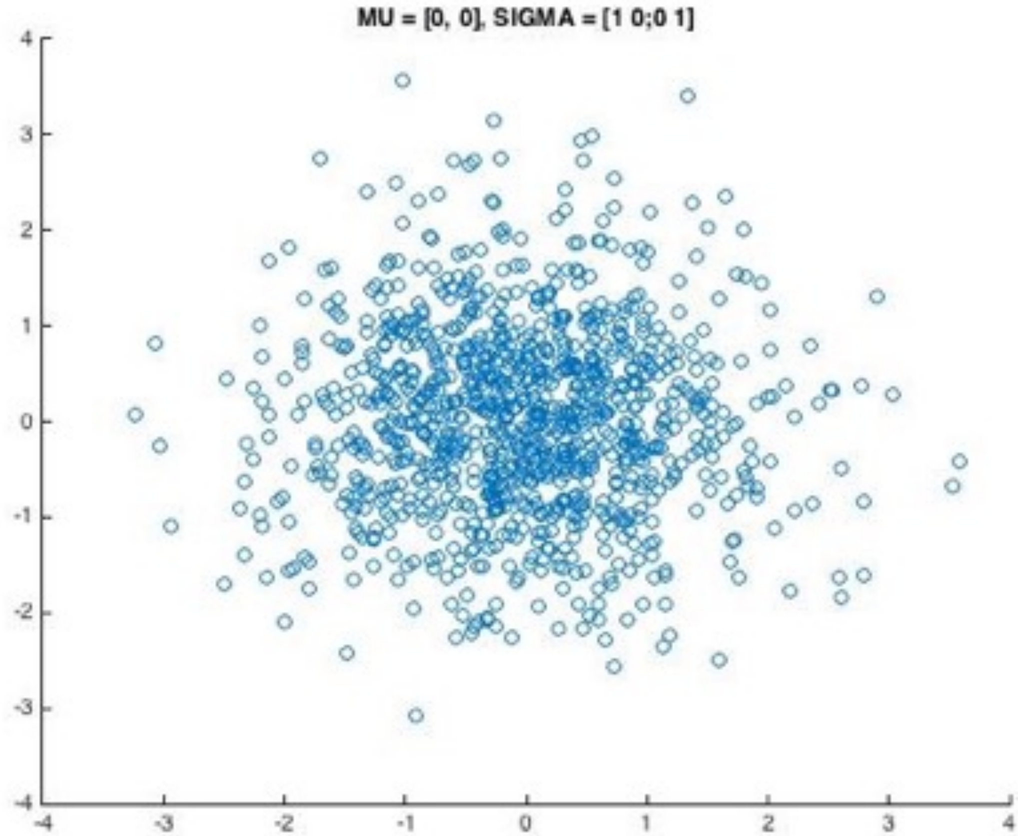
2、 The (x, y, z) on plane $Ax + by + Cz + D_2 = 0$

$$X = \frac{-AD_2}{A^2+B^2+C^2} \quad Y = \frac{-BD_2}{A^2+B^2+C^2} \quad Z = \frac{-CD_2}{A^2+B^2+C^2}$$

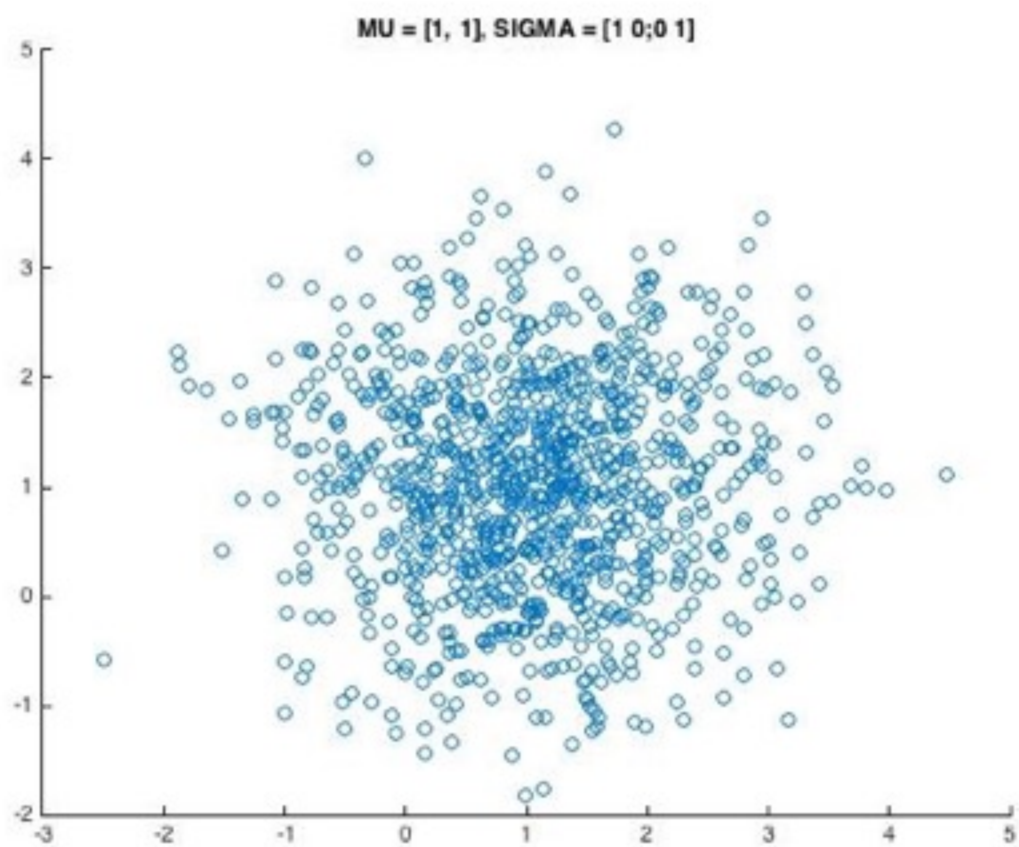
$$d = \frac{|Ax + By + Cz + D_1|}{\sqrt{A^2+B^2+C^2}} = \frac{|D_1 - D_2|}{\sqrt{A^2+B^2+C^2}},$$

therefore, the distance between parallel hyperplane $\begin{cases} w^T x + b_1 = 0 \\ w^T x + b_2 = 0 \end{cases}$

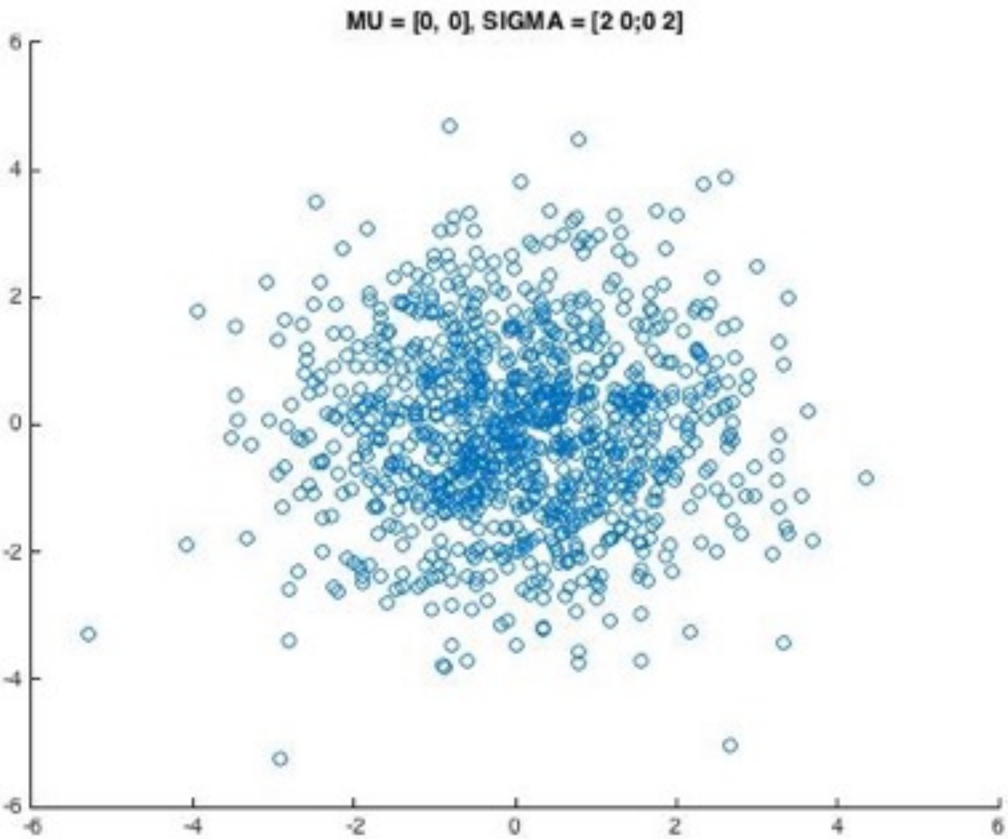
is $\frac{|b_2 - b_1|}{\|w\|_2}$ #

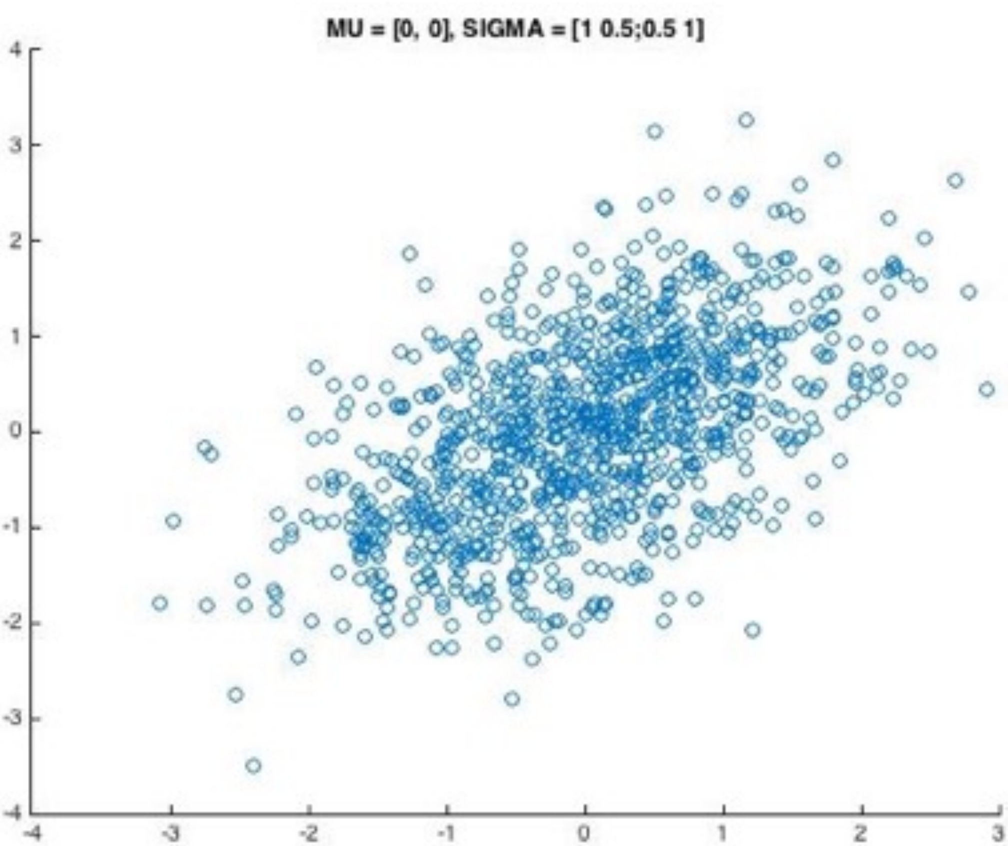


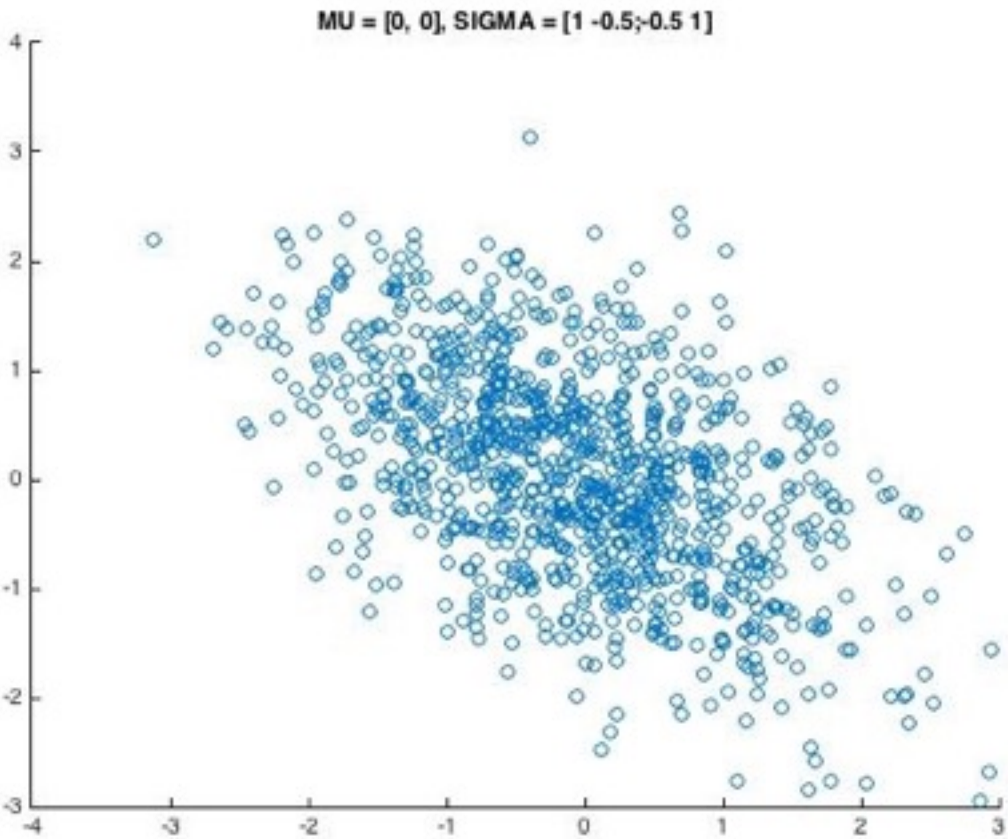
MU = [0, 0], SIGMA = [1 0;0 1]



MU = [1, 1], SIGMA = [1 0;0 1]


MU = [0, 0], SIGMA = [2 0;0 2]
6
4
2
0
-2
-4
-6
-6
-4
-2
0
2
4
6


MU = [0, 0], SIGMA = [1 0.5;0.5 1]
4
3
2
1
0
-1
-2
-3
-4
-4
-3
-2
-1
0
1
2
3


MU = [0, 0], SIGMA = [1 -0.5;-0.5 1]
4
3
2
1
0
-1
-2
-3
-4
-3
-2
-1
0
1
2
3