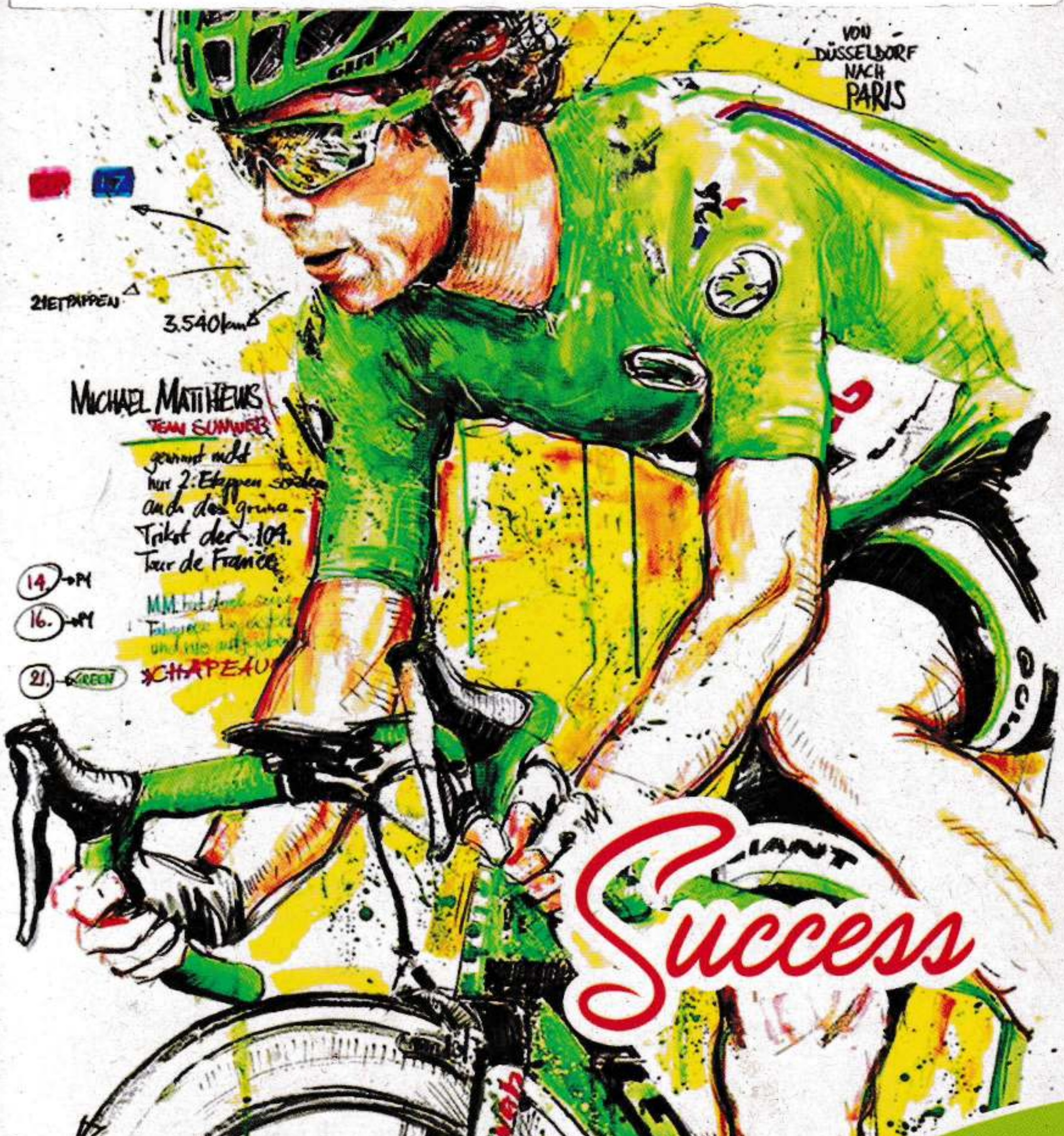
Introduction to Linear Algebra
- Gilbert Strang

18

Linear Algebra in
Probability and Statistics



Note Book

soorajss1729@gmail.com
+91-9400635788

INDEX

(18)

Name : SOORAJ. S. Subject :

Std. : Div. : Roll No. :

School / College :

S. No.	Date	Title	Page No.	Teacher's Sign / Remarks
		INTRODUCTION TO		
		LINEAR ALGEBRA		
		- Gilbert Strang, MIT		
		(5^{th} Edition)		

P.S 12.1

2. $\frac{1}{3}$ of all integers are divisible by 3 and $\frac{1}{7}$ of integers are divisible by 7. What fraction of integers will be divisible by 3 or 7 or Both ?

Ans: $$\frac{1}{3}+\frac{1}{7}-\frac{1}{3}\times\frac{1}{7}=\frac{1}{3}+\frac{1}{7}-\frac{1}{21}=\frac{9}{21}=\frac{3}{7}$$

3- Suppose you sample from the numbers 1 to 1000 with equal probabilities $1/1000$. What are the probabilities P_0 to P_9 that the last digit of your sample is $0, \dots, 9$? What is the expected mean m of that last digit ? What is its variance σ^2 ?

Ans: $$P_i = 100 \times \frac{1}{1000} = \frac{1}{10}$$

Expected mean of that last digit x :

$$m = E[x] = \sum_i P_i x_i = \frac{1}{10}\sum_{i=0}^{9} i = \frac{45}{10} = \underline{\underline{4.5}}$$

$$\sigma^2 = E[x^2] - (E[x])^2 = \frac{1}{10}\sum P_i x_i^2 - (4.5)^2$$
$$= \frac{1}{10}\sum_{i=0}^{9} i^2 - (4.5)^2 = \frac{285}{10} - (4.5)^2 = \underline{\underline{8.25}}$$

5. Sample again from 1 to 1000 with equal probabilities and let x be the 1st digit ($x = 1$ if the number is 15). What are the probabilities p_1 to p_9 (adding to 1) of $x = 1, \ldots, 9$? What are the mean & variance of x?

Ans:
$$p_1 = \frac{112}{1000}$$

$$p_2 = p_3 = \cdots = p_9 = \frac{111}{1000}$$

$$m = \sum_i p_i x_i = \frac{112}{1000} \times 1 + \frac{111}{1000}(2+3+\cdots+9)$$

$$= \frac{112 + 111(44)}{1000} = \frac{4996}{1000} = 4.996 \approx 5$$

which is close to $\frac{1}{9}(1+2+\cdots+9) = 5$

$$\sigma^2 = E[x^2] - m^2 = \frac{112}{1000}(1^2) + \frac{111}{1000}(2^2+\cdots+9^2) - m^2$$

$$\approx \frac{112 + 111(284)}{1000} - m^2 \approx \frac{31635}{1000} - 5^2 = \underline{6.635}$$

6. Suppose you have $N = 4$ samples $157, 312, 696, 602$ in problem 5. What are the 1st digits x_1 to x_4 of the squares? What is the sample mean μ? What is the sample variance S^2?

Ans: The 1st digits of $157, 312, 696, 602$ are $2, 9, 4, 3$.

The sample mean is:

$$\mu = \frac{1}{4}(2+9+4+3) = \frac{18}{4} = \underline{\underline{4.5}}$$

The sample variance with $N-1 = 4-1 = 3$ is:

$$S^2 = \frac{1}{3}\left[(-2.5)^2 + (4.5)^2 + (-0.5)^2 + (-1.5)^2\right]$$

$$= \frac{1}{3}(29)$$

7. $\sigma^2 = \sum P_i (x_i - m)^2 = \sum P_i (x_i^2 - 2m x_i + m^2)$

$= \sum P_i x_i^2 - 2m \sum P_i x_i + m^2 \sum P_i$

$= \sum P_i x_i^2 - 2m^2 + m^2 = \sum P_i x_i^2 - m^2$

$= E[x^2] - m^2$

8. If all 24 samples from a population produce the same age $x = 20$, what are the sample mean μ and sample variance S^2? What if $x = 20$ or 21, 12 times each?

Ans. $\mu = \frac{1}{N} \sum x_i = 20$

$S^2 = \frac{1}{N-1} \sum (x_i - \mu)^2 = 0$

PS (12.2)

1.
(b) The sum of N independent flips (0 or 1) is the count of heads after N tries. The rule for the variance of a sum gives $\sigma^2 =$ ______.

Ans: Independent flips $\Longrightarrow$ $N \times N$ covariance matrix is diagonal.

The diagonal entries are the variances $\sigma^2 = pq = p - p^2$ for each flip.

Overall variance of the sum from N flips is:

$$\cancel{Y = \begin{bmatrix} 1 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = [} \quad Y = \begin{bmatrix} 1 & \cdots & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_N \end{bmatrix} = \begin{bmatrix} x_1 + \cdots + x_N \end{bmatrix}$$

$$= AVA^T = \begin{bmatrix} 1 & \cdots & 1 \end{bmatrix} V \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix} = N\sigma^2 = N(p - p^2)$$

2. What's the covariance σ_{kl} b/w the results $x_1, \dots, x_n$ of exp. 3 and the results $y_1, \dots, y_n$ of exp 5 ?

Ans: $\sigma_{35} = \sum_{\text{all } i,j} \sum P_{ij} (x_i - m_3)(y_j - m_5)$

3. For $M = 3$ experiments, the variance-covariance matrix V will be 3×3. There will be a probability P_{ijk} that the 3 outputs are x_i and y_j and z_k. Write down a formula for the matrix V

Ans: $V = \sum_{\text{all}} \sum_{i,j,k} \sum P_{ijk} \, U U^T$

where $U = \begin{bmatrix} x_i - \bar{x} \\ y_j - \bar{y} \\ z_k - \bar{z} \end{bmatrix}$

4.

Ans:

6.

alts

4. What is the covariance matrix V for $M=3$ independent expts with means m_1, m_2, m_3 and variances $\sigma_1^2, \sigma_2^2, \sigma_3^2$?

Ans: $$V = \begin{bmatrix} \sigma_1^2 & 0 & 0 \\ 0 & \sigma_2^2 & 0 \\ 0 & 0 & \sigma_3^2 \end{bmatrix}$$

6. The $n \times n$ matrix P contains joint probabilities $P_{ij} = \text{Prob}\left(X = x_i \text{ and } Y = y_j\right)$

$$\text{Conditional Probability}\left(Y = y_j \mid X = x_i\right) = \frac{P_{ij}}{P_{i1} + \cdots + P_{in}} = \frac{P_{ij}}{P_i}$$

4. What is the covariance matrix V for M = 3 independent expts with means m_1, m_2, m_3 and variances $\sigma_1^2, \sigma_2^2, \sigma_3^2$?

Ans:
$$V = \begin{bmatrix} \sigma_1^2 & 0 & 0 \\ 0 & \sigma_2^2 & 0 \\ 0 & 0 & \sigma_3^2 \end{bmatrix}$$

6. The $n \times n$ matrix P contains joint probabilities $P_{ij} = \text{Prob}\left(X = x_i \text{ and } Y = y_j\right)$

$$\text{Conditional Probability}\left(Y = y_j \mid X = x_i\right) = \frac{P_{ij}}{P_{i1} + \cdots + P_{in}} = \frac{P_{ij}}{P_i}$$

12.3

1. Two measurements of the same variable x give 2 equations $x = b_1$ and $x = b_2$. Suppose the means are zero and the variances are σ_1^2 and σ_2^2, with independent errors: V is diagonal with entries σ_1^2 and σ_2^2. Write the 2 equations as $Ax = b$ (A is 2×1).

find this best estimate $\hat{x}$ based on b_1 and b_2:

$$\hat{x} = \frac{b_1/\sigma_1^2 + b_2/\sigma_2^2}{1/\sigma_1^2 + 1/\sigma_2^2} \; ; \; E[\hat{x}\hat{x}^T] = \left(\frac{1}{\sigma_1^2} + \frac{1}{\sigma_2^2}\right)^{-1}$$

Ans:
$$\left.\begin{array}{l} x = b_1 \\ x = b_2 \end{array}\right\} \begin{bmatrix} 1 \\ 1 \end{bmatrix} \begin{bmatrix} x \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \end{bmatrix} \iff Ax = b$$

The covariance matrix V is diagonal since the measurements are independent:

$$V = \begin{bmatrix} \sigma_1^2 & 0 \\ 0 & \sigma_2^2 \end{bmatrix}$$

The weighted least square solution is given by:

$$A^T V^{-1} A \hat{x} = A^T V^{-1} b$$

$$A^T V^{-1} A = \begin{bmatrix} 1 & 1 \end{bmatrix} \begin{bmatrix} 1/\sigma_1^2 & 0 \\ 0 & 1/\sigma_2^2 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \frac{1}{\sigma_1^2} + \frac{1}{\sigma_2^2}$$

$$A^T V^{-1} b = \begin{bmatrix} 1 & 1 \end{bmatrix} \begin{bmatrix} 1/\sigma_1^2 & 0 \\ 0 & 1/\sigma_2^2 \end{bmatrix} \begin{bmatrix} b_1 \\ b_2 \end{bmatrix} = \frac{b_1}{\sigma_1^2} + \frac{b_2}{\sigma_2^2}$$

$$\Longrightarrow \hat{x} = \frac{b_1/\sigma_1^2 + b_2/\sigma_2^2}{1/\sigma_1^2 + 1/\sigma_2^2}$$

The variance of that estimate $\hat{x}$ should be written as

$$E\left[(\hat{x}-x)(\hat{x}-x)^T\right] = \mathrm{Cov}\left[\hat{x}\right] = \mathrm{cov}\left[Lb\right]$$

$$= L\,\mathrm{cov}[b]\,L^T = (A^T V^{-1} A)^{-1} A^T V^{-1} V V^{-1} A (A^T V^{-1} A)^{-1}$$

$$= (A^T V^{-1} A)^{-1} = \left(\frac{1}{\sigma_1^2} + \frac{1}{\sigma_2^2}\right)^{-1}$$

2. (a) In problem 1, suppose the 2nd measurement b_2 becomes super-small and its variance $\sigma_2 \to 0$. What is the best estimate $\hat{x}$ when σ_2 reaches zero?

Ans: $$\hat{x} = \frac{b_1/\sigma_1^2 + b_2/\sigma_2^2}{1/\sigma_1^2 + 1/\sigma_2^2}$$

$\sigma_2 \to 0$: ~~Ans~~

$$\hat{x} = \frac{b_1/\sigma_1^2 + b_2/\sigma_2^2}{1/\sigma_1^2 + 1/\sigma_2^2} \Rightarrow \frac{b_1\sigma_2^2 + b_2\sigma_1^2}{\sigma_2^2 + \sigma_1^2}$$

xing by $\sigma_1^2\sigma_2^2$

$$\to \frac{b_2\sigma_1^2}{\sigma_1^2} = b_2$$

(b) The opposite case has $\sigma_2 \to \infty$ and no information in b_2. What is now the best estimate $\hat{x}$ based on b_1 and b_2?

Ans: $$\hat{x} = \frac{b_1/\sigma_1^2 + b_2/\sigma_2^2}{1/\sigma_1^2 + 1/\sigma_2^2} \to \frac{b_1/\sigma_1^2}{1/\sigma_1^2} = b_1$$

i.e., we are getting no information from the totally unreliable measurement $x = b_2$.

3. If x and y are independent with probabilities $P_1(x)$ and $P_2(y)$, then $P(x,y) = P_1(x)\, P_2(y)$.
By separating double integrals into products of single integrals ($-\infty$ to $+\infty$) show that

$$\iint P(x,y)\, dx\, dy = 1 \quad \text{and} \quad \iint (x+y)\, P(x,y)\, dx\, dy = m_1 + m_2$$

Ans:

$$P(x,y) = P_1(x)\, P_2(y)$$

$$\iint P(x,y)\, dx\, dy = \iint P(x) P(y)\, dx\, dy = \int P(x)\, dx \int P(y)\, dy$$

$$= 1 \cdot 1 = 1$$

$$\iint (x+y)\, P(x,y)\, dx\, dy = \iint x\, P(x) P(y)\, dx\, dy + \iint y\, P(x)\, P(y)\, dx\, dy$$

$$= \int x P(x)\, dx \int P(y)\, dy + \int P(x)\, dx \int y P(y)\, dy$$

$$= m_x \cdot 1 + 1 \cdot m_y = m_x + m_y$$

4. Continue problem 3 for independent x, y to show that $P(x,y) = P_1(x)\,P_2(y)$ has

$$\iint (x-m_1)^2\, P(x,y)\, dx\, dy = \sigma_1^2$$

$$\iint (x-m_1)(y-m_2)\, P(x,y)\, dx\, dy = 0$$

So the 2×2 covariance matrix V is diagonal and its entries are ________.

~~Ans~~:

Ans:
$$\iint (x-m_1)^2\, p(x,y)\, dx\, dy = \int (x-m_1)^2\, p(x)\, dx \int p(y)\, dy$$
$$= \sigma_1^2$$

$$\iint (x-m_1)(y-m_2)\, p(x,y)\, dx\, dy = \int (x-m_1)\, p(x)\, dx \int (y-m_2)\, p(y)\, dy$$
$$= (0)(0) = 0$$

5. Show that the inverse of a 2×2 covariance matrix V is

$$V^{-1} = \begin{bmatrix} \sigma_1^2 & \sigma_{12} \\ \sigma_{12} & \sigma_2^2 \end{bmatrix}^{-1} = \frac{1}{1-\rho^2} \begin{bmatrix} 1/\sigma_1^2 & -\rho/\sigma_1\sigma_2 \\ -\rho/\sigma_1\sigma_2 & 1/\sigma_2^2 \end{bmatrix}$$

with correlation $\rho = \dfrac{\sigma_{12}}{\sigma_1 \sigma_2}$

This produces the exponent $-(x-m)^T V^{-1} (x-m)$ in a 2-variable Gaussian.

Ans:

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

$$V^{-1} = \begin{bmatrix} \sigma_1^2 & \sigma_{12} \\ \sigma_{12} & \sigma_2^2 \end{bmatrix}^{-1} = \frac{1}{\sigma_1^2\sigma_2^2 - \sigma_{12}^2} \begin{bmatrix} \sigma_2^2 & -\sigma_{12} \\ -\sigma_{12} & \sigma_1^2 \end{bmatrix}$$

$$= \frac{1}{\sigma_1^2\sigma_2^2(1-\rho^2)} \begin{bmatrix} \sigma_2^2 & -\sigma_{12} \\ -\sigma_{12} & \sigma_1^2 \end{bmatrix}$$

$$\rho = \frac{\sigma_{12}^2}{\sigma_1\sigma_2}$$

$$= \frac{1}{1-\rho^2} \begin{bmatrix} 1/\sigma_1^2 & -\rho/\sigma_1\sigma_2 \\ -\rho/\sigma_1\sigma_2 & 1/\sigma_2^2 \end{bmatrix}$$

6.

Kalman filter

Success