1) Parameter estimation:
   i) A coin is tossed 10 times, with the results: H T T H H T H H H T. What is the maximum-likelihood estimate of its bias? If you want to make a Bayesian estimate of its parameter, how does the estimate change with every toss, starting from a prior of Beta(1,1)?
   ii) We are counting the number of buses that pass a given bus-stop every hour, which is [3 4 2 5 3 4 4 1 2 7 4]. This is considered to be following a Poisson Distribution. What is the probability that 6 buses will pass over the next 2 hours?
   iii) On every day, we record the maximum and minimum temperatures. We assume that they follow Gaussian distribution, but there is strong correlation among them. Calculate the parameters.
   [40, 22], [41, 24], [39, 20], [40, 20], [43 21], [45 25], [43 22], [42 22], [40, 20], [42, 23]

2) Conditional probability:
   There is a coin with bias p=0.7. There is a fair dice, and a loaded dice which produces only odd numbers (with equal probability). If the coin toss produces a Head, the fair dice is rolled, otherwise the loaded dice. The process is repeated twice, and the sum of the dice readings is 8. What is the probability that the fair dice was used both times?

3) Expectation:
   There are two 2D Gaussian distributions: one has parameters: mean [1 4], covariance [1 0; 0 1]. The other one has parameters: mean [2 2], covariance [1 2; 2 1]. A coin is tossed: if the result is Head we draw a sample from the first Gaussian, otherwise from the second.
   i) If the bias of the coin is known, how will you find the Expected value of the sampled vector?
   ii) If 10 observations sampled from the Gaussian distributions are known (but not which distribution), how can we estimate the bias of the coin?

4) We have four-hourly recordings of temperature for 7 days. Verify if the process is weakly stationery or not (with respect to mean, covariance, or correlation). Note that the estimated statistics need not be exactly equal even if the process is stationery.

	0:00AM	4:00AM	8:00AM	12noon	4:00PM	8:00PM
Day1	12	10	14	25	23	18
Day2	13	10	16	26	24	20
Day3	13	11	15	27	25	23
Day4	15	13	18	30	28	23
Day5	16	15	17	30	28	22
Day6	14	11	20	28	25	21
Day7	18	16	20	33	29	25

5) Once again, we use the same data as (4). I want to express this as an autoregressive process of order either 1 or 2. Frame the equations, and estimate the parameters. Can we use the same set of coefficients for all time-points, or do we need different coefficients?

6) There is a set of 10 spatial locations, whose latitude-longitude coefficents are provided. For each location, other locations within a 1degree distance may have influence. Consider a

spatial autoregressive process, where the coefficient of each influencing location s' on any location s is given by w/distance(s,s') [distance calculated in terms of lat-lon], where w is a constant. Estimate w, and the local component of each location.

Lat	Lon	Day 1	Day 2	Day 3	Day 4	Day 5
20	60	4	5	5	6	5
20	61	5	5	6	7	5
20	62	7	7	6	9	8
21	60	3	3	3	5	5
21	61	4	5	4	6	6
21	62	6	8	7	7	6
22	60	3	2	2	5	4
22	61	4	3	3	5	6
22	62	5	6	6	7	7

7) Consider a Gaussian Process with 0 mean and covariance function C(s1,s2) = exp(-0.5||s1-s2||) where ||s1-s2|| represent the distance between locations s1 and s2 (in terms of coordinates). Consider two locations whose distance is 2. On a given day, the measurement in one location is 5. What is the most likely measurement in the other location?

8) Consider a stationery spatial process, where the mean value of the variable at every location is 5, and the covariance function is defined as in the previous question (7). Given the values at 9 locations on 5 days (same table as 6), use Kriging to estimate the values at (21.5,61.5).

9) Consider a spatial variable X(s) which can be written as X(s) = m(s) + Y(s). m(s) is the fixed local component (mean), Y(s) is the global component which follows a Gaussian Process with mean function 0 and covariance function C(s,s') =1 if dist(s,s')<2, 0 otherwise. Consider 6 locations, whose lat-lon coordinates are given. In 1 of these locations, the local mean is not known, and in another location, X is not observed one day. Given the remaining observations of X and the mean, estimate the missing values.

Lat	Lon	Day 1	Day 2	Day 3	Day 4	Day 5	Local mean
20	60	4	5	5	6	5	5
20	61	5	5		7	5	6
20	62	7	7	6	9	8	7
21	60	3	3	3	5	5	
21	61	4	5	4	6	6	5
21	62	6	8	7	7	6	6

10) Consider rainfall X(s,t), where s is a location and t is the month of a year. X(s,t) may be decomposed as X(s,t)=a*m(s)+(1-a)*n(t)+ e(s,t) where m is location-specific constant, n is month-specific constant, a is constant and e(s,t) is Gaussian noise N(0,0.1). X(s,t) is measured every month of every year. If a and m(s1) is known, using the observations of X, estimate m and n.

11) Consider a spatio-temporal dataset for 20 years. Calculate the climatology for each location, and the anomaly value of each observation. Calculate the 90^{th} quantile of each location. Calculate the return periods of the climatology value for each location. Check if the positive anomalies and 90^{th} quantile events are temporally coherent, i.e. is $p(Y(t)>0|Y(t-1)>0) > p(Y(t)>0)$? Similarly, are such events spatially coherent, i.e. $p(Y(s,t)>0|Y(s',t)>0) > p(Y(s',t)>0)$, where (s,s') are neighbors?

	T1	T2	T3	T4	T5	T6	T7	T8	T9	T10
S1	45	48	49	44	38	39	46	40	42	48
S2	54	56	59	50	42	46	52	48	52	56
S3	60	58	67	61	52	56	68	50	58	62
	T11	T12	T13	T14	T15	T16	T17	T18	T19	T20
S1	46	48	42	39	32	38	58	52	47	48
S2	55	57	51	48	40	48	68	59	51	55
S3	62	61	57	53	52	58	66	64	58	66
	T21	T22	T23	T24	T25	T26	T27	T28	T29	T30

12) Given the following spatio-temporal dataset with missing values, explore the different approaches for imputation, such as spatial/temporal averaging, and matrix factorization (assuming that the data matrix is 'approximately low-rank')? [(s1, s2), (s2, s3), (s3, s4) are neighbours]

	T1	T2	T3	T4	T5	T6
S1	20	16	18	19	X	17
S2	25	X	23	24	21	21
S3	18	30	X	20	29	28
S4	15	36	26	19	34	X

13) Consider a spatio-temporal dataset with missing values. Suppose the measurements at each location follows a Gaussian distribution, whose parameters are known. Also, it is known that on any day, the difference between the measurements at any two locations follows a Gaussian distribution with mean 0 and variance proportional to the distance between the locations. Using this information, predict the missing values using Bayesian analysis.

Locations	Gaussian	T1	T2	T3	T4	T5
12N,70E	(10,5)	12	16	7	X	17
13N,72E	(12,2)	13	X	9	10	X
14N,74E	(8,4)	11	11	X	X	15
15N,76E	(15,5)	X	11	10	11	19

14) A set of annual observations are available at some locations. Suppose each of them follow a Gamma distribution with known shape and scale parameters. Define anomaly events as those which have a return period of 10 years or more with respect to those distributions. Identify the anomaly events. Also identify such bulk anomaly years where in the same year the values at 2 or more locations have a return period of at least 5 years.

	(shape, scale)	Y1	Y2	Y3	Y4	Y5	Y6	Y7	Y8
S1	5,10	64	78	31	30	39	57	34	120
S2	8,10	71	127	112	69	54	75	82	96
S3	2,10	16	5	25	21	23	59	1	4
S4	10,10	110	69	107	77	141	98	107	101

15) We consider the annual maximum rainfall (block maxima) at different locations. We know that they follow a GEV distribution, with scale (σ=25) and shape (ξ=0). The location (μ) parameters at the different locations are known to follow a Gaussian process with mean function 5 (constant) and covariance function C(s,s') = exp(-0.5*||s-s'||). If μ(s1) is known, estimate μ at the other locations.

Lat	Lon	Y1max	Y2max	Y3max	Y4max	Y5max	Y6max	Y7max	Y8max
5N	10E	6	38	42	-8	13	10	25	31
6N	10E	38	0	30	28	-9	-13	15	86
5N	11E	2	20	-6	35	-4	14	30	58
6N	11E	85	18	-12	-11	-3	49	-3	45

16) Consider a set of 7 locations, where a variable is measured each day. Construct a correlation network between these locations using a suitable cutoff. The climatological value and variance of the variable at each location is also provided. Identify the extreme events at each location and construct an event synchronization network, with lag of 1 day.

S1	S2	S3	S4	S5	S6	S7
12	14	21	15	6	3	15
3	2	4	3	3	1	1

	S1	S2	S3	S4	S5	S6	S7
T1	16	11	18	12	5	0	17
T2	11	12	22	9	8	2	15
T3	13	17	21	11	12	3	16
T4	15	20	19	12	13	2	18
T5	18	22	24	16	10	-3	19
T6	15	19	17	15	7	-2	16
T7	10	17	15	12	4	-2	9
T8	8	11	13	12	1	0	10
T9	7	10	9	14	-2	2	13
T10	11	13	12	17	1	3	17
T11	15	16	15	20	5	4	18

17) Consider a linear system with a 2D state vector x(t), from which 1D observations y(t) are available. The system equations are: [x(t+1) = Ax(t) + g, y(t) = Bx(t) + h] where g, h are Gaussian Noises. The initial state and the relevant matrices are known. At each time-step, predict the next observation when you receive the current observation.

A = [1 -1; -1 2]; B = [2 -1] x(1) = [0 0]; Var(g)= [0.5 0; 0 1]; Var(h)=2;

Y = [-0.07 0.9 2.82 11.73 23.0];

18) Consider a linear system with 2D state vector x(t) from which 2D observations y(t) are available. The system equations are: [x(t+1) = Ax(t) + g, y(t) = Bx(t) + h] where g, h are Gaussian Noises. Given the system state measurements and observations at each time-step, estimate the observation and state transition matrices A and B as well as the noise parameters.

X1	0	0.32	0.28	2.64	6.40	19.16
X2	0	-0.59	-1.62	-5.35	-12.19	-30.58
Y	-0.07	0.9	2.82	11.73	23.0	66.98

Solutions :

③

i) $N_1\left(\begin{bmatrix}1\\4\end{bmatrix}, \begin{bmatrix}1&0\\0&1\end{bmatrix}\right)$ $N_2\left(\begin{bmatrix}2\\2\end{bmatrix}, \begin{bmatrix}1&2\\2&1\end{bmatrix}\right)$

$z_i \sim \text{Ber}(p) \to 1, 2$

$\mathbb{E}(x_i)$ need to find

$x_{z=1} \sim N_1$

$x_{z=2} \sim N_2$

$\mathbb{E}(x_i) = \int x \, PDF \, dx$

$= p\begin{bmatrix}1\\4\end{bmatrix} + (1-p)\begin{bmatrix}2\\2\end{bmatrix}$

$= \begin{bmatrix}p + 2(1-p)\\ 4p + 2(1-p)\end{bmatrix} = \begin{bmatrix}2-p\\ 2p+2\end{bmatrix}$

ii) $x_1\ x_2 \cdots x_{10}$ (z unknown)

Now how can we find p ??

$P(x) = p(x, z=1) + p(x, z=2)$

Let $x_i = \begin{bmatrix}2\\3\end{bmatrix}$ $= p(z=1)\,p(x|z=1) + p(z=2)\,p(x|z=2)$

$= p\,N_1\left(\begin{bmatrix}2\\3\end{bmatrix}\right) + (1-p)N_2\left(\begin{bmatrix}2\\3\end{bmatrix}\right)$

$= pe^{-1} + (1-p)ae^{-b}$

∴ Doing MLE similarly for $x_2, \ldots, x_{10}$

$$P(x_1, x_2, \ldots, x_{10}) = \prod P(x_i) = \mathcal{L}$$

$$\therefore \frac{\partial \mathcal{L}}{\partial p} = 0 \quad \Leftrightarrow \quad = 0 \text{ gives us bias of coin 'p'}$$

④ if weekly stationarity / not (wrt mean / cov / corr)

a) $\mu(0\text{AM}) = 15$
$\mu(4\text{AM}) = 13$
$\mu(8\text{AM}) = 14$
$\mu(12\text{PM}) = 25$
$\mu(4\text{PM}) = \vdots$
$\mu(8\text{PM}) = !$

} Not Mean Statio.

b) if $\text{Cov}(X_{t_1}, X_{t_2}) = f(|t_1 - t_2|)$

ie. $\text{Cov}(Y_0, Y_4) = \text{Cov}(X_4, X_8) = \text{Cov}(X_8, X_{12})$

$\text{Cov} = E(AB) - E(A)E(B)$

$(A, B) \to x_0, x_4 = \boxed{\ldots}$

$\text{Cov}(X_4, X_8) = \ldots$

} Now Check

Now if $\Delta t = 4$ matches, check $\Delta t = 8$, $\Delta t = 12$ etc.

If all holds only then Cov. Station.

⑤ 1st Order Auto-Reg. Process

$$X_t = a X_{t-1} + b$$

2nd Order: $X_t = a_1 X_{t-1} + a_2 X_{t-2} + b$

Do Least Square !!!

$$E = \sum_{d=1}^{z} e_d^2$$

1st order $\begin{cases} \frac{\partial E}{\partial a} = 0 & \frac{\partial E}{\partial b} = 0 \end{cases}$

2nd order $\begin{cases} \frac{\partial E}{\partial a_1} = 0 & \frac{\partial E}{\partial a_2} = 0 & \frac{\partial E}{\partial b} = 0 \end{cases}$

(t, t+1)

For each ~~day~~ Same param/diff

Check which case error more

Choose other

⑥ $$X_S = \underbrace{\mu_S}_{\text{local}} + \underbrace{W_1 X_1 + W_2 X_2 + \dots + W_S X_S}_{\text{Global}}$$

$X_2 = \mu_2 + W X_1 + W X_3$ ← (Now $W_i = 0$ if dist > 1, $= 1$ else)

Write data and eqns and apply Least square

$\mu = 0$

⑦ $$Cov(S_1, S_2) = \exp\left(-\frac{1}{2}\|S_1 - S_2\|\right)$$ $\|S_1 - S_2\| = 2$

Now $X_{S_1} = 5$ $X_{S_2} = ?$

$$C(X_1, X_2) = e^{-\frac{1}{2}^2} = e^{-1}$$

$X \sim GP(\mu, C)$

$$\begin{bmatrix} X_1 \\ X_2 \end{bmatrix} \sim \mathcal{N}\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} Var(X_1) & e^{-1} \\ e^{-1} & Var(X_2) \end{bmatrix}\right)$$

$Var(X_1) = Cov(X_1, X_1) = 1 = Var(X_2)$

$$\therefore \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \sim N\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \underbrace{\begin{bmatrix} 1 & e^{-1} \\ e^{-1} & 1 \end{bmatrix}}_{C} \right)$$

$$P\left(\begin{bmatrix} s \\ x \end{bmatrix}\right) = \frac{1}{\begin{vmatrix} 1 & e^{-1} \\ e^{-1} & 1 \end{vmatrix}} \exp\left[-\frac{1}{2} (s-x)^T C^{-1} (s-x) \right]$$

MLE : $\frac{\partial P}{\partial x} = 0$, this gives $\underline{x}$

↓ i.e. most likely value in position s_2

(a) $X(s) = m(s) + Y(s) \cdot m(s)$

$Y(s) \sim GP(0, C(s,s'))$ $\quad C(s,s') = 1$ if $\|s,s'\| < 2$

a) Day 3 at (20', 61') unknown = x

b) For (21', 60') mean not known [ignore now]

a) $\therefore Y(s) = X(s) - m(s)$

$$Y = \begin{bmatrix} 0 \\ x-c \\ -1 \\ -1 \\ 1 \end{bmatrix} \sim N\left(\begin{bmatrix} 2 \\ \vdots \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 & 0 & \cdots \\ 1 & \ddots & \ddots \\ \ddots & \ddots & 1 \\ \cdots & 1 & 1 \end{bmatrix} \right)$$

PDF

$$p\left(\begin{bmatrix} 0 \\ x-6 \\ -1 \\ -1 \\ 1 \end{bmatrix}\right) = \frac{1}{|\Sigma|} \quad ----$$

$\frac{\partial p}{\partial x} = 0 \longrightarrow$ Find missing value x.

b) Easy way: $\mu_4 = \frac{\text{Simple sum}}{\text{sample}} = \frac{\text{Dev}_1 + \text{Dev}_2 \cdots}{S}$

(But as not LLN not perfect)

Better Way: Convert Table of X into Y

$Y = X - \mu$ Take gaussian PDF MLE

$$\begin{pmatrix} 0 \\ \vdots \\ 3-\mu \\ \vdots \end{pmatrix} = \text{---}.$$

Similar way $\boxed{\frac{\partial p}{\partial \mu} = 0}$

(10) $X(s,t) = a\,m(s) + (1-a)\,n(t) + e(s,t)$

↳ every month of year.

$a \rightarrow$ known

$m(s_1) \rightarrow$ known

$e(s,t) \sim N(0, 0.1)$

Find all m

Find all n

$X(s_1,t_1) = a\,m(s_1) + (1-a)\,n(t_1) + e(s_1,t_1)$

$X(s_1,t_2) =$ - - - -

$x_1 = \underbrace{a\,m_1}_{c_1} + (1-a)\,n_j + e_1$

$x_2 = e_1 + (1-a)\,n_j + e_2$

⋮ ⋮ ⋮

$\bar{X} = N e_1 + N(1-a)\,n_j + \Sigma e$

N ↳ no. of years

assuming large N $\Sigma e \approx 0$

MLE on Σe find n_{jan}

Similarly find $\mu_{Feb}, \mu_{Mar}, \ldots$ etc.

(11) → Climatology = Mean value (Long term)

→ Anomaly = Obs − Climatology

→ 90th percentile

Sort values in order then

⇒ choose 'value' which is 18 values less than it

20 values

$\therefore$ 90th per $= \frac{20 \times 90}{100}$

$= 18$ values

→ Return period for $x = 40$

No. of obs $\leq 40 = 6$

$\therefore p = \frac{6}{20} = 0.3$

Return period $= \frac{1}{0.3} =$ **3.33 years**

→ if +ve anomaly and 90th percentile event are temporally corr.

$$P(\underbrace{Y(t)>0}_{\text{anomalies already calcu in 1st step}}) = \frac{9}{20} \text{ (let)}$$

$$P(Y(t)>0 \mid Y(t-1)>0) = \frac{6}{9} = \frac{2}{3}$$

$$\therefore \frac{9}{20} < \frac{2}{3} \Rightarrow \text{+ve anamoly events are tempo. corre.}$$

90^{th} quantile event $= \frac{3}{20}$

$$P(90^{th} \text{per}(t) \mid 90^{th} \text{per}(t-1)) = \frac{1}{3}$$

$$\frac{3}{20} < \frac{1}{3} \Rightarrow \text{Also tempo. corr.}$$

(12) Missing Value prediction using Low Rank Matrix Factoriz

For S_2 → better temporal avg. or spatial/tempo. avg.

For S_4 / S_3 → better spacial avg in clusters

Best Low Rank Matrix Factorization

$X = A \cdot B$ $\quad X_{4\times 6}$

$4\times 1 \quad 1\times 6$

or

$4\times 2 \quad 2\times 6$

(13) Measurement at each loc ~ Gaussian (param. known)

On same day

Δ(measurm of loc) ~ Gaussian $(0, \sigma^2)$

$\sigma^2 \propto$ ||dist||

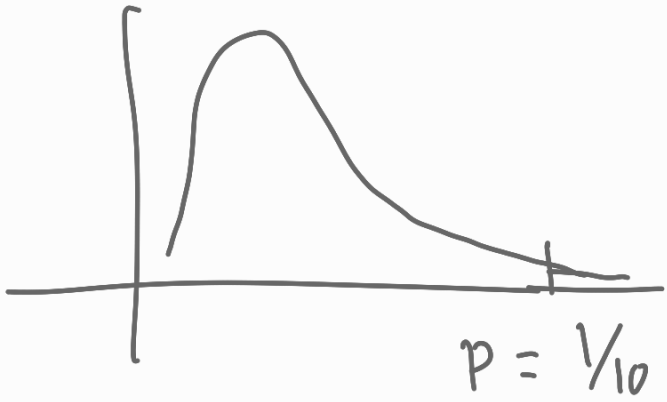
$$P(x) = \frac{1}{\sqrt{2\pi}\cdot 2} \exp\left[-\frac{1}{2\sigma^2}(x-\mu)^2\right]$$

$\sigma = 2$

(14) Gamma distri

Return period = 10 years

Anamoly



(15) GEV(Loca, Shape, scale)
μ ξ σ