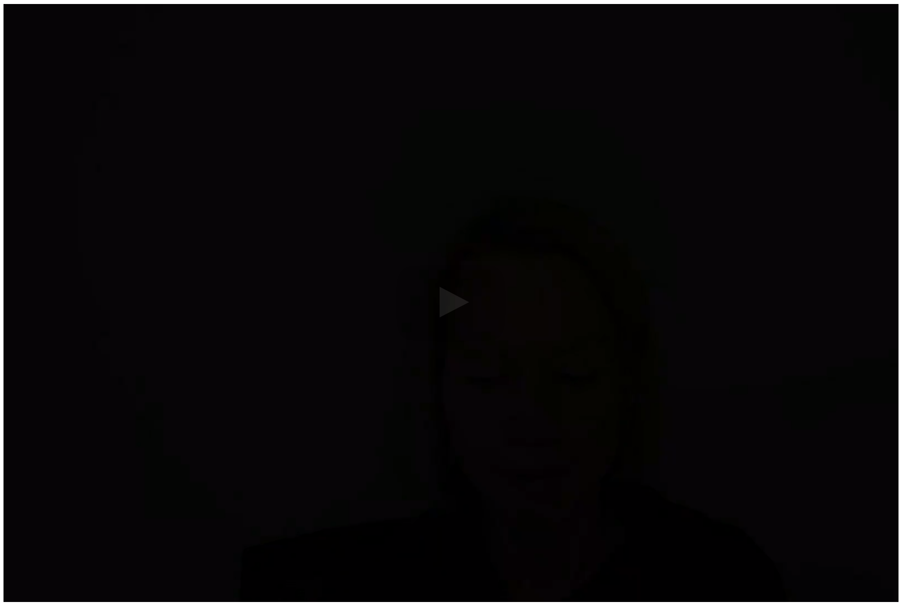
Exercises due May 19, 2021 19:59 EDT

Spatial Correlation



16:32 / 16:32 1.50x

And that'll be the key to our next steps.

So all these intuitions sort of are leading us

towards this much more powerful model that,

at the end of the day, will allow us to approximate very

flexible patterns in data.

End of transcript. Skip to the start.

Video

Download video file

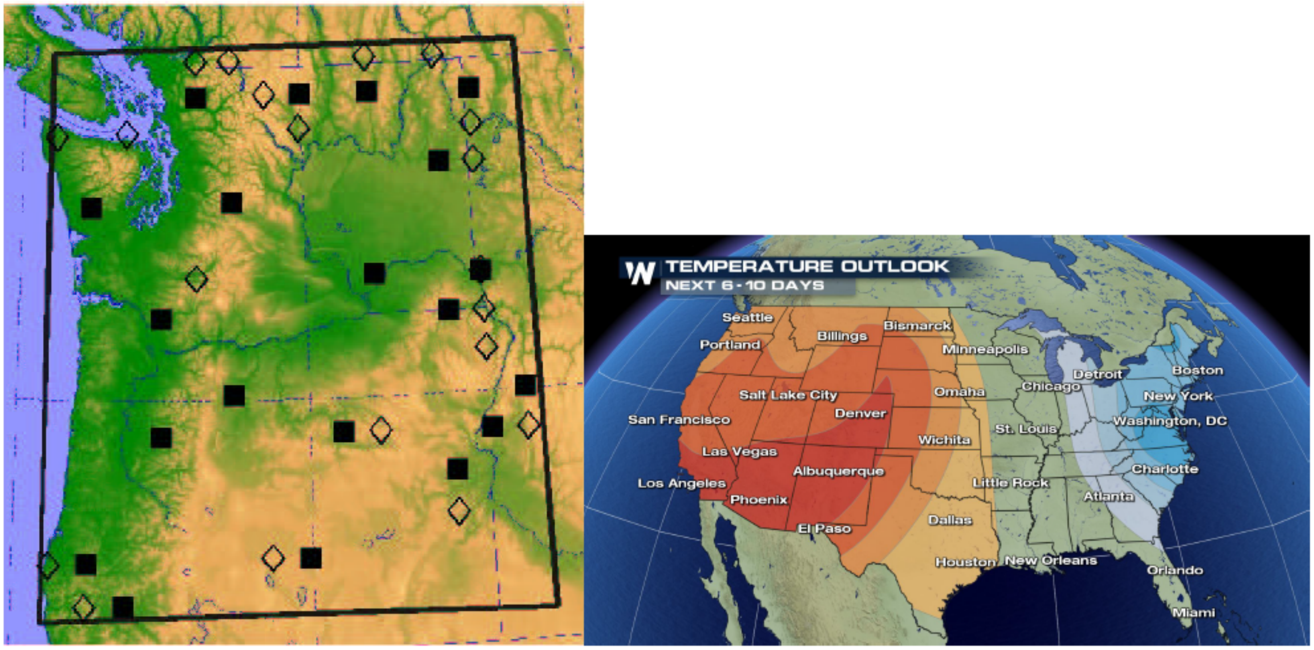
Transcripts

Download SubRip (.srt) file

Download Text (.txt) file

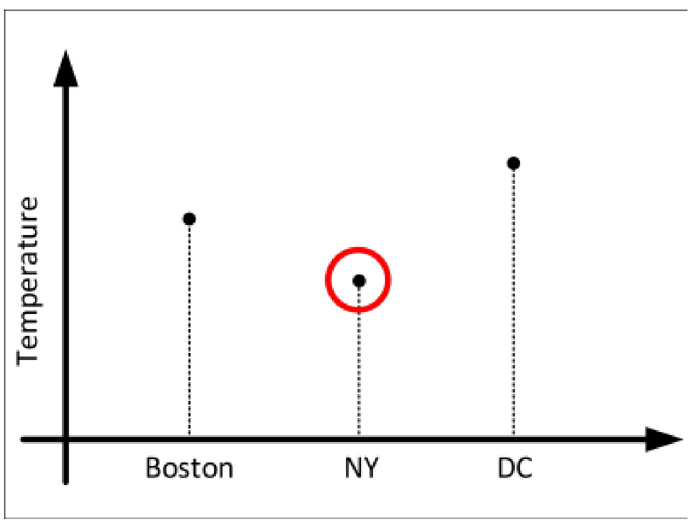
One central intuition is that two measurements will be correlated as a function of their distance, i.e., the distance between the points where the measurements are taken.

For example: we have a number of locations, as shown as black squares in the below figure, for which we have temperature measurements. One might be interested in estimating the temperature at some intermediate locations, the diamonds in the figure.



27: Some examples of spatial correlation. The image on the right shows other examples of temperature correlations.

Let us say we have access to temperature measurements on three cities: Boston, New York, and DC, as shown in the figure below. Given that we know that the cities are close to each other we expect that the temperatures will be correlated. So, even if we do not know the temperature in New York, we can create some form of prediction based on the temperatures Boston and DC. We can try to make a good guess knowing that they are close to each other, and their measurements are correlated.



28: Temperature values at three different cities.

This is essentially a regression problem when we want to interpolate the measurement based on our observations. This is what we are going to study next. We will develop a method to predict the unknown temperature with their corresponding certainty measure of the prediction by assuming that we have access to the correlation between the variables and the temperatures behave as Gaussian random variables.

We will start by reviewing a number of definitions and concepts that will be useful later for our estimation of correlations in space.

Definitions

Definition 4.1 [Normal Distribution] A normal distribution is a continuous probability distribution for a real-valued random variable. The unidimensional probability density function is defined as

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}}\exp\left(-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right),$$

where the parameter μ is the mean or expectation of the distribution, and σ is the standard deviation. The variance is σ^2. The French school calls this distribution the Gaussian distribution.

Gaussian Distribution

1 point possible (graded)
What is the median of a Normal distribution with mean μ and standard deviation σ?

- ◯ $\frac{\mu^2}{\sigma}$
- ◯ $\frac{\mu}{\sigma}$
- ◯ μ
- ◯ $\sigma\mu$

Submit You have used 0 of 2 attempts

Standard Normal Distribution

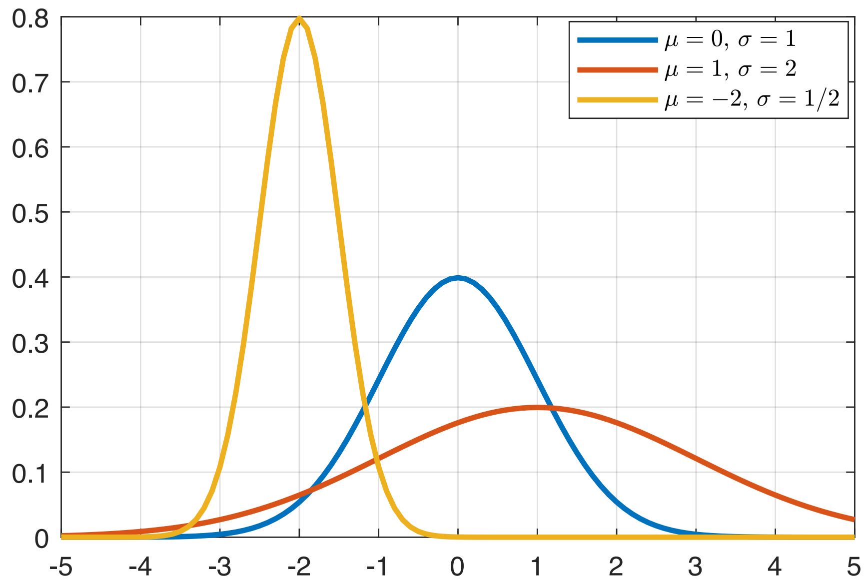
1 point possible (graded)
The standard Normal distribution is defined as the Normal distribution with mean $\mu = 0$ and standard deviation $\sigma = 1$. What is the maximum value attained by the probability density function of a standard Normal distribution?

- ◯ 1
- ◯ $\sqrt{2\pi}$

- $\frac{1}{\sqrt{2\pi}}$
- 2π

Submit You have used 0 of 2 attempts

The below figure shows three examples of a Gaussian Distribution with various parameter values for μ and σ.



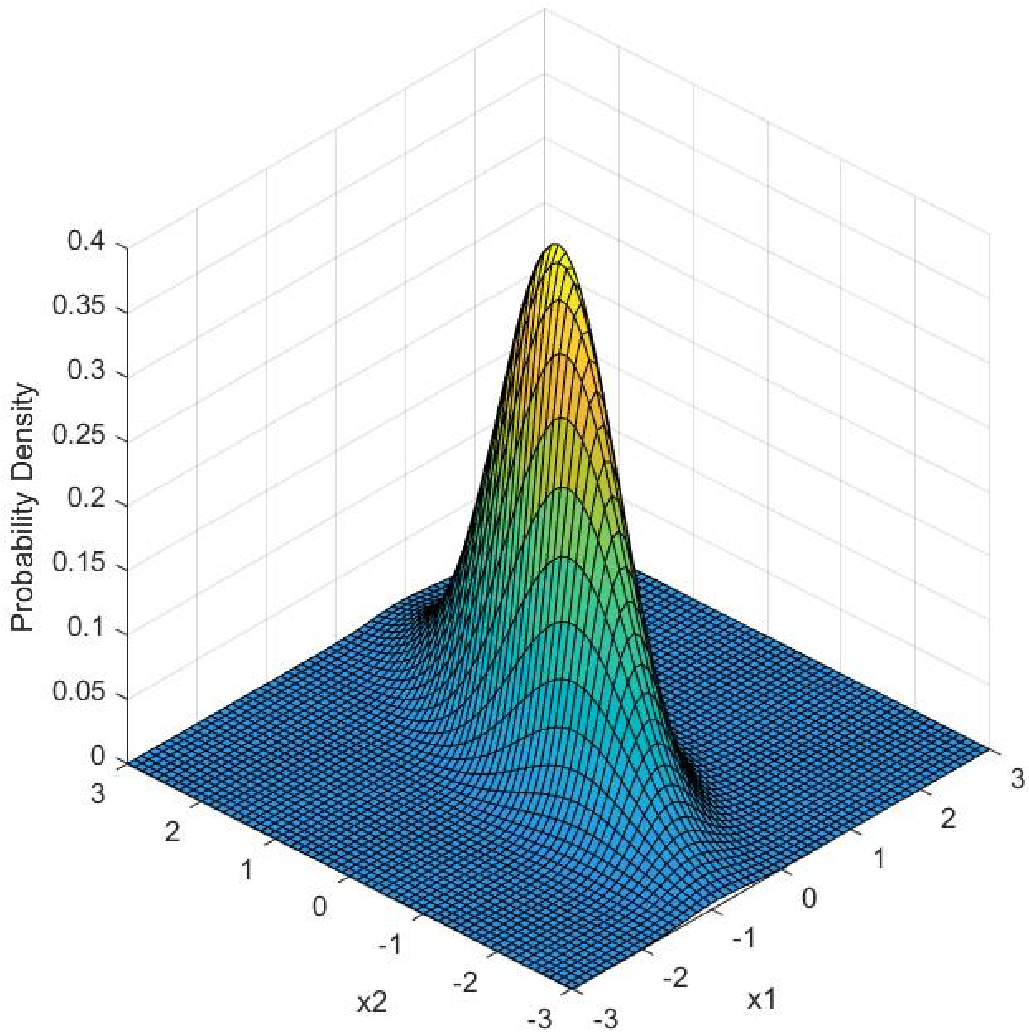
29: Three examples of Normal distributions, with various values of μ and σ.

Definition 4.2 The probability density function of a Multivariate Gaussian random variable is defined as

$$p\left(x \mid \mu, \Sigma\right) = \frac{1}{(2\pi)^{n/2}|\Sigma|^{1/2}}\exp\left(-\frac{1}{2}(x-\mu)^{\intercal}\Sigma^{-1}\left(x-\mu\right)\right),$$

where $\mu = \mathbb{E}\left[X\right]$, and $|\Sigma|$ denotes the determinant of the covariance matrix Σ.

An example of this multivariate distribution is shown as a 3D plot below.



30: An example of the pdf for a multivariate normal distribution.

Definition 4.3 [Covariance] **Covariance** quantifies the joint variation between two random variables. Formally, the covariance between two real-valued random variables X and Y, both with finite second moments, is defined as

$$\text{Cov}\,(X, Y) = \mathbb{E}\,[(X - \mu_X)\,(y - \mu_Y)] = \mathbb{E}\,[XY] - \mathbb{E}\,[X]\,\mathbb{E}\,[Y]\,.$$

Some other notations for the covariance are σ_{XY} or $\sigma\,(X, Y)$.

Definition 4.4 [Correlation] **Correlation** is defined as any statistical relation between two random variables. There are many ways we can measure correlation. The Pearson correlation coefficient between two random variables X and Y, with expected values $\mathbb{E}\,[X] = \mu_X$ and $\mathbb{E}\,[Y] = \mu_Y$, and standard deviations σ_X and σ_Y, is defined as

$$\rho_{XY} = corr(X,Y) = \frac{\text{Cov}(X,Y)}{\sigma_X,\sigma_Y} = \frac{\mathbb{E}[(X-\mu_X)(y-\mu_Y)]}{\sigma_X\sigma_Y}.$$

Below are some properties of the covariance operator:

- Cov $(X, X) = \sigma_X^2$ also known as variance.
- Cov $(aX + bY, cW + dV) = ac*$Cov $(X, W) + ad*$Cov $(X, V) + bc*$Cov $(Y, W) + bd*$Cov (Y, V). (This mean that the covariance operator is bilinear.)

.

Properties of Covariance

1 point possible (graded)
Given two random variables X and Y with covariance Cov $(X, Y) = 0.5$, Cov $(X, X) = 1$ and Cov $(Y, Y) = 2$, define a new random variable $Z = 2X + Y$. What is the value of Cov (Z, Z)?

Submit You have used 0 of 4 attempts

Definition 4.5 Two random variables X and Y are independent if and only if, for every x and x, the events $\{X \leq x\}$ and $\{Y \leq y\}$ are independent. Recall that two events A and B are independent if and only if their joint probability equals the product of their probabilities, i.e., $P(A \cap B) = P(A)P(B)$. For random variables, this is equivalent to

$$F_{X,Y}(x,y) = F_X(x)F_Y(y),$$

where $F_X(x)$ and $F_Y(y)$ are the cumulative functions of the random variables X and Y.

Properties of Random Variables

1 point possible (graded)
If X and Y are independent random variables, and both are defined on a finite sample space $\Omega = z_1, \cdots, z_n$, then $\mathbb{E}[XY] = \mathbb{E}[X]\mathbb{E}[Y]$, as

$$\begin{aligned}\mathbb{E}[XY] &= \sum_{i,j=1}^{n} z_i z_j P(X = z_i, Y = z_j) \\ &= \sum_{i,j=1}^{n} z_i z_j P(X = z_i)\, P(Y = z_j) \\ &= \left(\sum_{i=1}^{n} P(X = z_i)\, z_i\right)\left(\sum_{j=1}^{n} z_j P(Y = z_j)\right).\end{aligned}$$

What is $\mathbb{E}[XY]$ when X and Y represent two unbiased independent six sided dice?

Submit You have used 0 of 4 attempts

Properties of Random Variables

1 point possible (graded)

Assume you are given two unbiased six sided dice modeled as random variables X_1 and X_2 respectively. However, die 2 depends on the output of die 1. If the output of the die 1 is even, then die 2 can only output odd numbers. Note there is no condition if die 1 is odd. What is $\mathbb{E}[X_1 X_2]$.

Submit You have used 0 of 4 attempts

Properties of Random Variables

2 points possible (graded)

Define two random variables as $X \sim U[-1, 1]$ and $Y = X^2$:

- Are they independent?

  - Yes
  - No

- Are they correlated?

- ◯ Yes
- ◯ No

Submit You have used 0 of 1 attempt

Properties of Random Variables

1 point possible (graded)

Compute the correlation of random variables X and Y with the following joint distribution.

$P(X=x,Y=y)$	$y=-1$	$y=0$	$y=+1$
$x=0$	0	1/3	0
$x=1$	1/3	0	1/3

Submit You have used 0 of 4 attempts

Properties of Covariance Matrix

1 point possible (graded)

Is the following matrix a valid covariance matrix?

$$\begin{bmatrix} 11 & -3 & 7 & 5 \\ -3 & 11 & 5 & 7 \\ 7 & 5 & 11 & -3 \\ 5 & 7 & -3 & 11 \end{bmatrix}$$

- ◯ Yes, because it is symmetric and positive semi-definite

- ◯ Yes, because it is symmetric
- ◯ No, because it is not symmetric or it is not positive semi-definite
- ◯ No, because it is not symmetric

Submit You have used 0 of 1 attempt

Properties of Random Variables

1 point possible (graded)
Can a non-symmetric matrix be a covariance matrix?

- ◯ Yes
- ◯ No

Submit You have used 0 of 1 attempt

Discussion

Hide Discussion

Topic: Module 5: Environmental Data and Gaussian Processes:Environmental Data and Gaussian Processes / 4. Spatial Correlation

Add a Post

Show all posts by recent activity

? Properties of Random Variables <The 2 die problem> 2
Approach: 1. When X1 is even, assume X2 can only be odd and calculated the probability as p(X1=x, X2=y) = p(X2=y|X1=x...

[staff] formula typo on the page 24 of the slide 1
The formula displayed on the slide is kindly wrong. The power of 2*pi in the denominator is marked as -d/2. However I be...

[STAFF] Can you double check answer to 'Properties of Random Variables' 7
I thought it was straightforward but the answer I get is marked wrong. Cannot think of a mistake I made either.