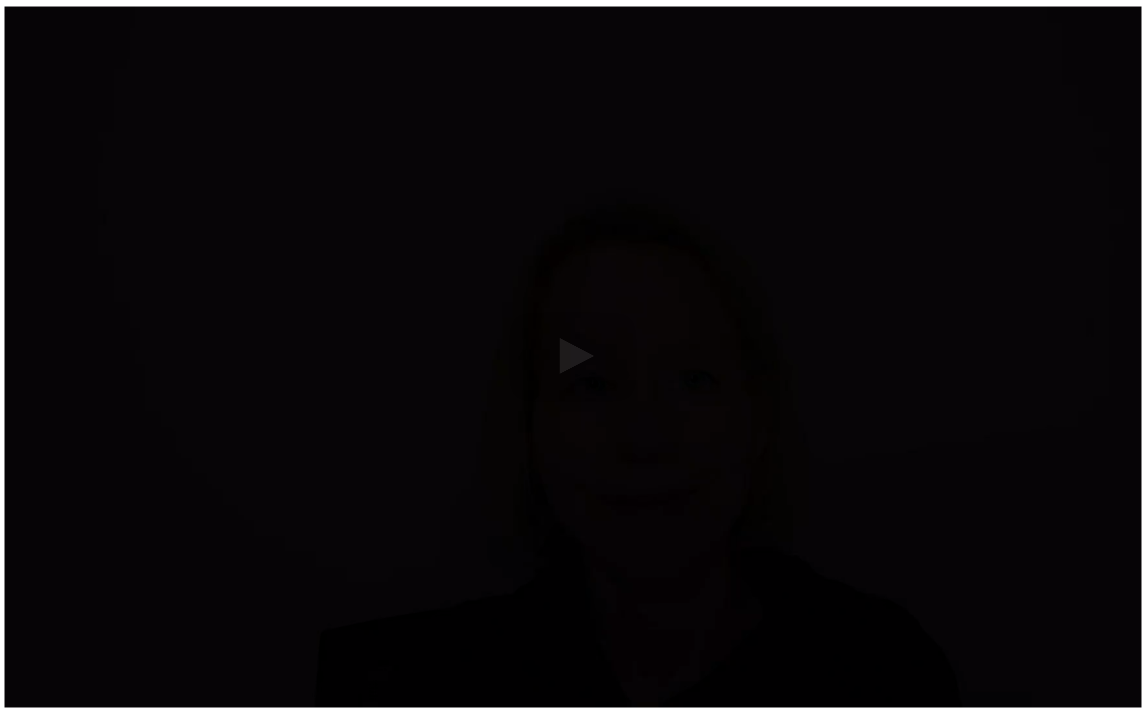
6. Conditional Probability for Multiple Variables

Bookmark this page

Conditional Probability for Multiple Variables



And this is a topic we actually will

continue in the next lecture.

So so far, we looked at, OK, we have a multivariate Gaussian,

and that seems to actually capture this fairly well.

And we get, magically, the smooth behavior for function

out.

And in the next lecture, we will see

how we can actually model this and tune this

and target it to our needs.

13:58 / 13:58 1.50x

Video
Download video file

Transcripts
Download SubRip (.srt) file
Download Text (.txt) file

Conditional Distribution of Multivariate Gaussian Random Variables

Following the same arguments as before we can construct a general definition of the conditional distribution for joint Normal random variables of any dimension. Consider a N-dimensional multivariate Gaussian random variable partitioned as follows

$$\mathbf{X} = \begin{bmatrix} \mathbf{X}_1 & \in \mathbb{R}^d \\ \mathbf{X}_2 & \in \mathbb{R}^{N-d} \end{bmatrix}$$

where

$$\mu_{\mathbf{X}} = \begin{bmatrix} \mu_1 & \in \mathbb{R}^d \\ \mu_2 & \in \mathbb{R}^{N-d} \end{bmatrix}$$

and

$$\mathbf{\Sigma} = \begin{bmatrix} \mathbf{\Sigma}_{11} \in \mathbb{R}^{d\times d} & \mathbf{\Sigma}_{12} \in \mathbb{R}^{d\times(N-d)} \\ \mathbf{\Sigma}_{21} \in \mathbb{R}^{(N-d)\times d} & \mathbf{\Sigma}_{22} \in \mathbb{R}^{(N-d)\times(N-d)} \end{bmatrix}$$

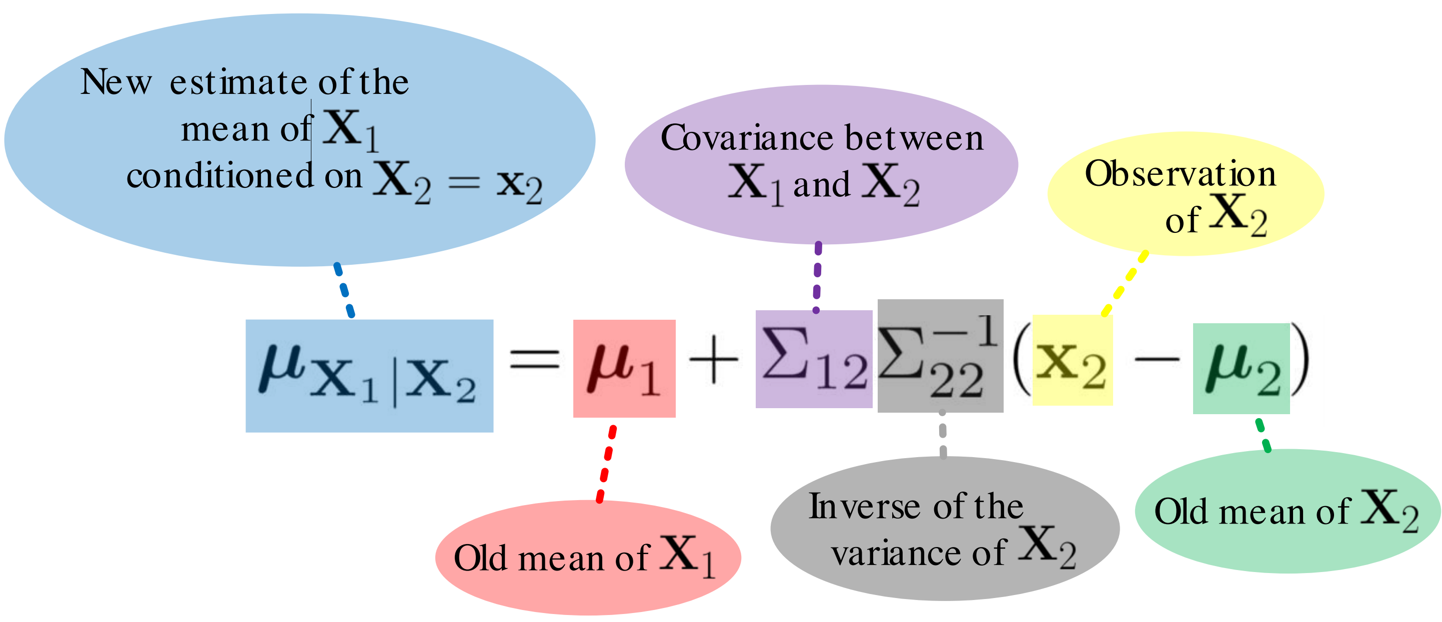
The distribution of the random variable $\mathbf{X}_1$ conditioned on $\mathbf{X}_2 = \mathbf{x}_2$ is a Gaussian distribution with

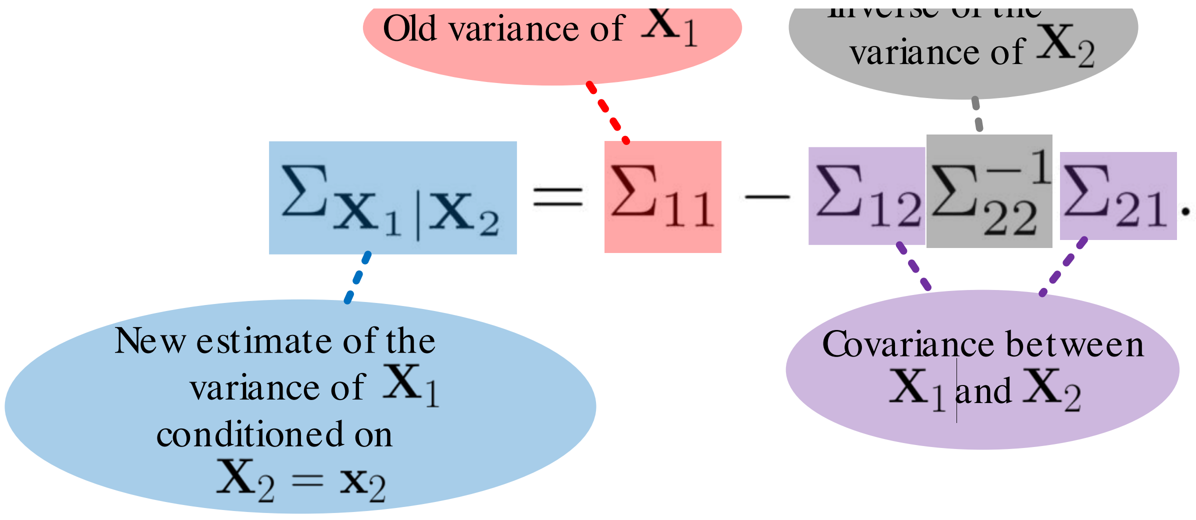
$$\mu_{\mathbf{X}_1|\mathbf{X}_2} = \mu_1 + \Sigma_{12}\Sigma_{22}^{-1}(\mathbf{x}_2 - \mu_2) \quad \text{and} \quad \Sigma_{\mathbf{X}_1|\mathbf{X}_2} = \Sigma_{11} - \Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21}. \tag{7.4}$$

A summary of the symbols used, their dimension, their meaning, and availability:

Symbol	Dimension	Meaning	Availability
$\mathbf{X}_1$	$\mathbb{R}^d$	Unobserved random variables, e.g., temperatures at cities for which we do not have measurements but we would like to estimate.	We know $\mathbb{E}(\mathbf{X}_1)$ and $var(\mathbf{X}_1)$
$\mathbf{X}_2$	$\mathbb{R}^{N-d}$	Observed random variables, e.g., temperatures at cities for which we have measurements.	We know $\mathbb{E}(\mathbf{X}_2)$ and $var(\mathbf{X}_2)$
μ_1	$\mathbb{R}^d$	Mean of the unobserved random variables $\mathbf{X}_1$	We know $\mathbb{E}(\mathbf{X}_1) = \mu_1$
μ_2	$\mathbb{R}^{(}N-d)$	Mean of the observed random variables $\mathbf{X}_2$	We know $\mathbb{E}(\mathbf{X}_2) = \mu_2$
$\mathbf{x}_1$	$\mathbb{R}^d$	Realizations of the random variable $\mathbf{X}_1$	We do NOT have access to this.
$\mathbf{x}_2$	$\mathbb{R}^{N-d}$	Realizations of the random variable $\mathbf{X}_2$	We do have access to this.
Σ_{11}	$\mathbb{R}^{d\times d}$	Variance of the random variables $\mathbf{X}_1$	We have access to this $var(\mathbf{X}_1)$.
Σ_{12}	$\mathbb{R}^{d\times(N-d)}$	Covariance of the random variables $\mathbf{X}_1$ with $\mathbf{X}_2$	We have access to this Cov $(\mathbf{X}_1, \mathbf{X}_2)$.
Σ_{21}	$\mathbb{R}^{(N-d)\times d}$	Covariance of the random variables $\mathbf{X}_2$ with $\mathbf{X}_1$	We have access to this Cov $(\mathbf{X}_2, \mathbf{X}_1)$.
Σ_{22}	$\mathbb{R}^{(N-d)\times(N-d)}$	Variance of the random variables $\mathbf{X}_2$	We have access to this $var(\mathbf{X}_2)$.
$\mu_{\mathbf{X}_1\|\mathbf{X}_2}$	$\mathbb{R}^d$	New estimate of the mean of $\mathbf{X}_1$ given the observations $\mathbf{x}_2$ of the random variable $\mathbf{X}_2$	We compute this.
$\Sigma_{\mathbf{X}_1\|\mathbf{X}_2}$	$\mathbb{R}^{d\times d}$	New estimate of the variance of $\mathbf{X}_1$ given the observations $\mathbf{x}_2$ of the random variable $\mathbf{X}_2$	We compute this.

The below figure shows a visual guide on the meaning of all the terms in the multivariate Gaussian parameter update rule.





32: The elements of mutivariate Gaussian conditional distributions.

As a final note, observe that the process described in this section builds the conditional distribution of the random variable $\mathbf{X}_1$ given $\mathbf{X}_2$. Since we have assumed both they form a jointly Gaussian random variable, their conditional distribution is also Gaussian with the new parameters given in the condition equations above.

We will conclude this chapter with a formal definition of a Gaussian process.

Definition 6.1 A Gaussian process is a collection of indexed random variables, (indexed for example by time or space), such that every finite subset of those random variables are jointly Gaussian.

For more information

You may wish to consult the book: C. E. Rasmussen & C. K. I. Williams, Gaussian Processes for Machine Learning, the MIT Press, 2006, ISBN 026218253X.

http://www.GaussianProcess.org/gpml

http://www.gaussianprocess.org/gpml/chapters/RW.pdf

Hide

Discussion

Hide Discussion

Topic: Module 5: Environmental Data and Gaussian Processes:Environmental Data and Gaussian Processes / 6. Conditional Probability for Multiple Variables

Add a Post

Show all posts by recent activity

STAFF What if instead we fit a regression or polynomial

Previous Next

edX

About
Affiliates
edX for Business
Open edX
Careers

News

Legal

Terms of Service & Honor Code
Privacy Policy
Accessibility Policy
Trademark Policy
Sitemap

Connect

Blog
Contact Us
Help Center
Media Kit
Donate