ECE 417 Lecture 4: Multivariate Gaussians

Mark Hasegawa-Johnson

9/7/2017

Content

- Vector of i.i.d. Gaussians
- Vector of Gaussians that are independent, but not identically distributed
- Some facts about linear algebra
- The Mahalanobis form of the multivariate Gaussian
- The Mahalanobis form for Gaussians that are not independent
- More facts about linear algebra
- More facts about ellipses

Vector of I.I.D. Gaussian Variables

Suppose we have a frame containing N samples from a Gaussian white noise process, $x_1, \ldots, x_N$. Let's stack them up to make a vector:

$$\vec{x} = \begin{bmatrix} x_1 \\ \vdots \\ x_N \end{bmatrix}$$

This whole frame is random. In fact, we could say that $\vec{x}$ is a sample value for a Gaussian random vector called $\vec{X}$, whose elements are $X_1, \ldots, X_N$:

$$\vec{X} = \begin{bmatrix} X_1 \\ \vdots \\ X_N \end{bmatrix}$$

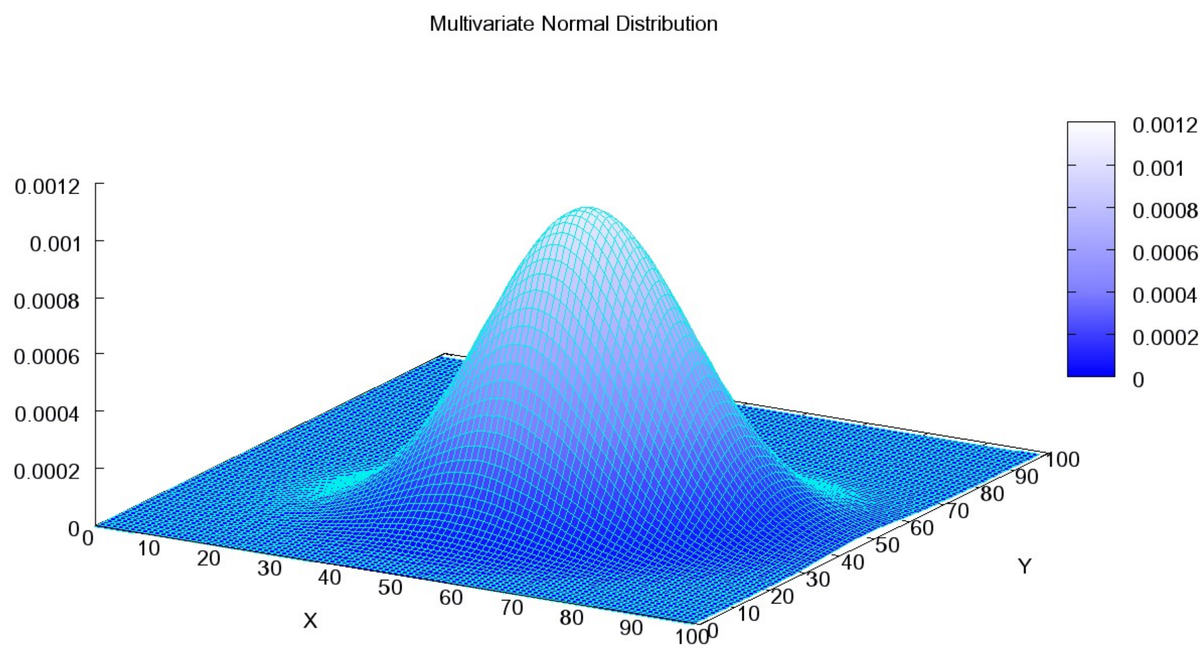
Vector of I.I.D. Gaussian Variables

Suppose that the N samples are i.i.d., each one has the same mean, μ, and the same variance, σ^2. Then the pdf of this random vector is

$$f_{\vec{X}}(\vec{x}) = \mathcal{N}(\vec{x}; \vec{\mu}, \sigma^2 I) = \prod_{n=1}^{N} \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2}\left(\frac{x_n - \mu}{\sigma}\right)^2}$$

Vector of I.I.D. Gaussian Variables

For example, here's an example from Wikipedia with mean of 50 and standard deviation of about 12.



Attribution: Piotrg,
https://commons.wikimedia.org/wiki/File:Multivariate_Gaussian.png

Independent Gaussians that aren't identically distributed

Suppose that the N samples are independent Gaussians that aren't identically distributed, i.e., X_d has mean μ_d and variance ${\sigma_d}^2$. The pdf of X_d is

$$f_{X_d}(x_d) = \mathcal{N}(x_d; \mu_d, {\sigma_d}^2) = \frac{1}{\sqrt{2\pi{\sigma_d}^2}} e^{-\frac{1}{2}\left(\frac{x_d-\mu_d}{\sigma_d}\right)^2}$$

The pdf of this random vector is

$$f_{\vec{X}}(\vec{x}) = \mathcal{N}(\vec{x}; \vec{\mu}, \Sigma) = \prod_{d=1}^{D} \frac{1}{\sqrt{2\pi{\sigma_d}^2}} e^{-\frac{1}{2}\left(\frac{x_d-\mu_d}{\sigma_d}\right)^2}$$

Independent Gaussians that aren't identically distributed

Another useful form is:

$$\prod_{d=1}^{D} \frac{1}{\sqrt{2\pi{\sigma_d}^2}} e^{-\frac{1}{2}\left(\frac{x_d-\mu_d}{\sigma_d}\right)^2} = \frac{1}{(2\pi)^{D/2}\prod_{d=1}^{D}\sigma_d} e^{-\frac{1}{2}\sum_{d=1}^{D}\left(\frac{x_d-\mu_d}{\sigma_d}\right)^2}$$

Example

Suppose that $\mu_1 = 1, \mu_2 = -1, {\sigma_1}^2 = 1, {\sigma_2}^2 = 4$. Then

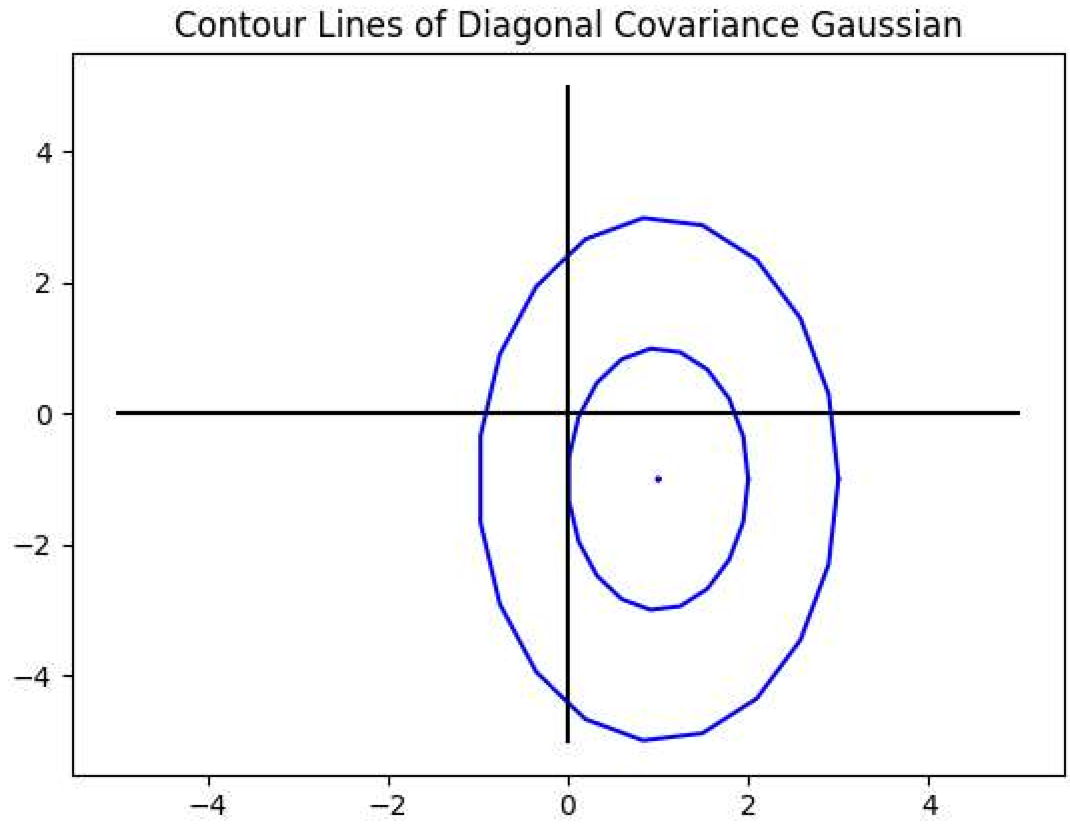
$$f_{\vec{X}}(\vec{x}) = \prod_{d=1}^{2} \frac{1}{\sqrt{2\pi{\sigma_d}^2}} e^{-\frac{1}{2}\left(\frac{x_d-\mu_d}{\sigma_d}\right)^2} = \frac{1}{4\pi} e^{-\frac{1}{2}\left(\left(\frac{x_1-1}{1}\right)^2+\left(\frac{x_2+1}{2}\right)^2\right)}$$

The pdf has its maximum value, $f_{\vec{X}}(\vec{x}) = \frac{1}{4\pi}$, at $\vec{x} = \vec{\mu} = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$.

It drops to $\frac{1}{4\pi\sqrt{e}}$ at $\vec{x} = \begin{bmatrix} \mu_1 \pm \sigma_1 \\ \mu_2 \end{bmatrix}$ and at $\vec{x} = \begin{bmatrix} \mu_1 \\ \mu_2 \pm \sigma_2 \end{bmatrix}$.

It drops to $\frac{1}{4\pi e^2}$ at $\vec{x} = \begin{bmatrix} \mu_1 \pm 2\sigma_1 \\ \mu_2 \end{bmatrix}$ and at $\vec{x} = \begin{bmatrix} \mu_1 \\ \mu_2 \pm 2\sigma_2 \end{bmatrix}$.

Example


Contour Lines of Diagonal Covariance Gaussian
4
2
0
−2
−4
−4
−2
0
2
4

Example

- OK, things are going to get even more complicated, so let’s remember what that means.
- It means that there are two Gaussian random variables, x1 and x2.
- X1 is Gaussian with an average value of 1, and a variance of 1.
- X2 is Gaussian with an average value of -1, and a variance of 4.
- Got it? OK. Let’s keep going.

Facts about linear algebra #1: determinant of a diagonal matrix

Suppose that Σ is a diagonal matrix, with variances on the diagonal:

$$\Sigma = \begin{bmatrix} {\sigma_1}^2 & 0 & 0 \\ 0 & {\sigma_2}^2 & \dots \\ 0 & \dots & {\sigma_D}^2 \end{bmatrix}$$

Then the determinant is

$$|\Sigma| = \prod_{d=1}^{D} {\sigma_d}^2$$

So we can write the Gaussian pdf as

$$\frac{1}{(2\pi)^{D/2}|\Sigma|^{1/2}} e^{-\frac{1}{2}\sum_{d=1}^{D}\left(\frac{x_d-\mu_d}{\sigma_d}\right)^2} = \frac{1}{|2\pi\Sigma|^{1/2}} e^{-\frac{1}{2}\sum_{d=1}^{D}\left(\frac{x_d-\mu_d}{\sigma_d}\right)^2}$$

Facts about linear algebra #2: inner product

Suppose that

$$\vec{x} = \begin{bmatrix} x_1 \\ \vdots \\ x_D \end{bmatrix} \text{ and } \vec{\mu} = \begin{bmatrix} \mu_1 \\ \vdots \\ \mu_D \end{bmatrix}$$

Then

$$(\vec{x} - \vec{\mu})^T (\vec{x} - \vec{\mu}) = (x_1 - \mu_1)^2 + \cdots + (x_D - \mu_D)^2$$

Facts about linear algebra #3: inverse of a diagonal matrix

Suppose that Σ is a diagonal matrix, with variances on the diagonal:

$$\Sigma = \begin{bmatrix} {\sigma_1}^2 & 0 & 0 \\ 0 & {\sigma_2}^2 & \dots \\ 0 & \dots & {\sigma_D}^2 \end{bmatrix}$$

Then its inverse, Σ^{-1}, is

$$\Sigma^{-1} = \begin{bmatrix} \frac{1}{{\sigma_1}^2} & 0 & 0 \\ 0 & \frac{1}{{\sigma_2}^2} & \dots \\ 0 & \dots & \frac{1}{{\sigma_D}^2} \end{bmatrix}$$

Facts about linear algebra #4: squared Mahalanobis distance with a diagonal covariance matrix

Suppose that all of the things on the previous slides are true.

Then the squared Mahalanobis distance is

$$d_\Sigma^2(\vec{x},\vec{\mu}) = (\vec{x}-\vec{\mu})^T\Sigma^{-1}(\vec{x}-\vec{\mu}) =$$

$$[x_1-\mu_1,\dots,x_D-\mu_D]\begin{bmatrix} \frac{1}{{\sigma_1}^2} & 0 & 0 \\ 0 & \frac{1}{{\sigma_2}^2} & \dots \\ 0 & \dots & \frac{1}{{\sigma_D}^2} \end{bmatrix}\begin{bmatrix} x_1-\mu_1 \\ \vdots \\ x_D-\mu_D \end{bmatrix}$$

$$= \frac{(x_1-\mu_1)^2}{{\sigma_1}^2} + \cdots + \frac{(x_D-\mu_D)^2}{{\sigma_D}^2}$$

Mahalanobis form of the multivariate Gaussian, independent dimensions

So we can write the multivariate Gaussian as

$$f_{\vec{X}}(\vec{x}) = \mathcal{N}(\vec{x}; \vec{\mu}, \Sigma) = \frac{1}{|2\pi\Sigma|^{1/2}} e^{-\frac{1}{2}(\vec{x}-\vec{\mu})^T \Sigma^{-1} (\vec{x}-\vec{\mu})}$$

$$f_{\vec{X}}(\vec{x}) = \mathcal{N}(\vec{x}; \vec{\mu}, \Sigma) = \frac{1}{|2\pi\Sigma|^{1/2}} e^{-\frac{1}{2} d_{\Sigma}^2 (\vec{x}-\vec{\mu})}$$

Facts about ellipses

The formula

$$1 = (\vec{x} - \vec{\mu})^T \Sigma^{-1} (\vec{x} - \vec{\mu})$$

… or equivalently

$$1 = \frac{(x_1 - \mu_1)^2}{{\sigma_1}^2} + \cdots + \frac{(x_D - \mu_D)^2}{{\sigma_D}^2}$$

… is the formula for an ellipsoid (an ellipse in two dimensions; a football shaped object in three dimensions; etc.). The ellipse is centered at the point $\vec{\mu}$, and it has a volume proportional to $|\Sigma|$. (In 2D the area of an ellipse is $\pi|\Sigma|^{1/2}$, in 3D it's $\frac{4}{3}\pi|\Sigma|^{1/2}$, etc.)

Facts about ellipses

$$c = (\vec{x} - \vec{\mu})^T \Sigma^{-1} (\vec{x} - \vec{\mu})$$

… is equivalent to

$$f_{\vec{X}}(\vec{x}) = \frac{1}{|2\pi\Sigma|^{1/2}} e^{-\frac{1}{2}c}$$

Therefore the contour plot of a Gaussian pdf --- the curves of constant $f_{\vec{X}}(\vec{x})$ --- are ellipses. If Σ is diagonal, the main axes of the ellipse are parallel to the x_1, x_2, etc. axes. If Σ is NOT diagonal, the main axes of the ellipse are tilted.

Mahalanobis form of the multivariate Gaussian, dependent dimensions

If the dimensions are dependent, and jointly Gaussian, then we can still write the multivariate Gaussian as

$$f_{\vec{X}}(\vec{x}) = \mathcal{N}(\vec{x}; \vec{\mu}, \Sigma) = \frac{1}{|2\pi\Sigma|^{1/2}} e^{-\frac{1}{2}(\vec{x}-\vec{\mu})^T \Sigma^{-1} (\vec{x}-\vec{\mu})}$$

Example

Suppose that x_1 and x_2 are linearly correlated Gaussians with means 1 and -1, respectively, and with variances 1 and 4, and covariance 1.

$$\vec{\mu} = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$$

Remember the definitions of variance and covariance:

$$\sigma_1{}^2 = E[(x_1 - \mu_1)^2] = 1$$
$$\sigma_2{}^2 = E[(x_2 - \mu_2)^2] = 4$$
$$\sigma_{12} = \sigma_{21} = E[(x_1 - \mu_1)(x_2 - \mu_2)] = 1$$

$$\Sigma = \begin{bmatrix} 1 & 1 \\ 1 & 4 \end{bmatrix}$$

Determinant and inverse of a 2x2 matrix

You should know the determinant and inverse of a 2x2 matrix. If

$$\Sigma = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$$

Then $|\Sigma| = ad - bc$ and

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

You should be able to verify the inverse, for yourself, by multiplying $\Sigma\Sigma^{-1}$ and discovering that the result is the identity matrix.

Example

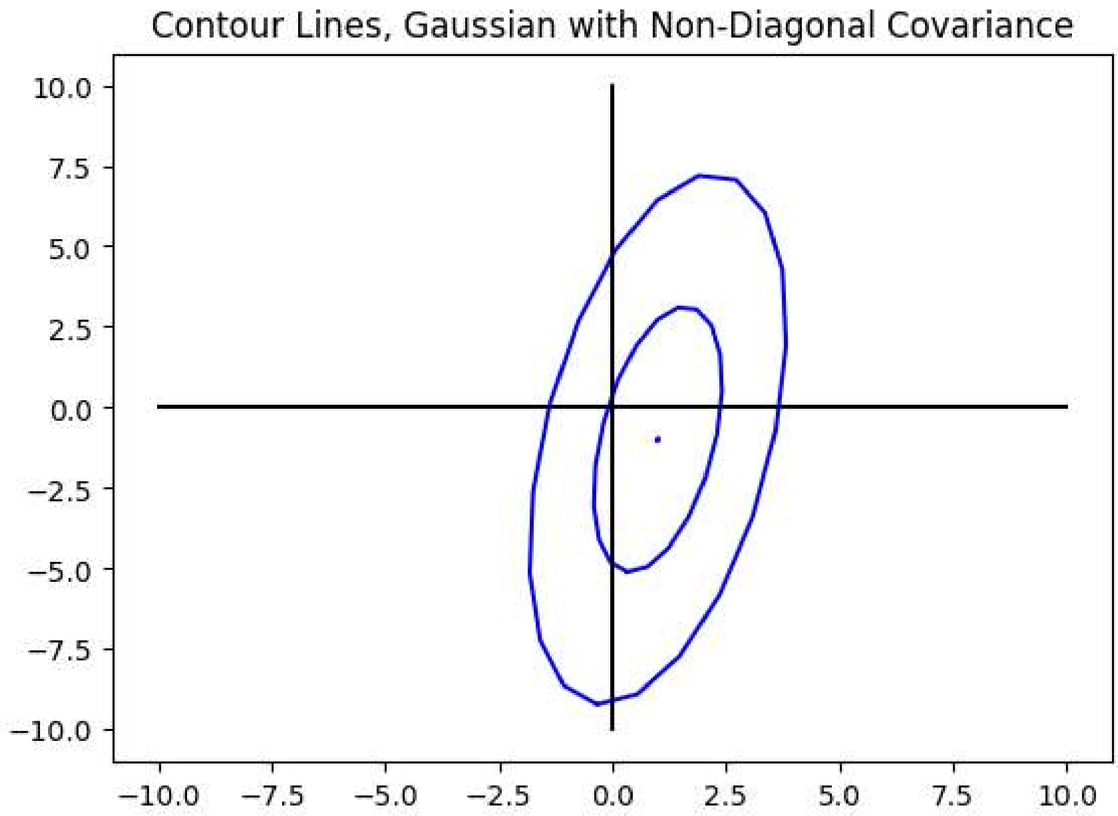
Therefore the contour lines of this Gaussian are ellipses centered at

$$\vec{\mu} = \begin{bmatrix} 1 \\ -1 \end{bmatrix}.$$

The contour lines are ellipses that satisfy this equation. Each different value of c gives a different ellipse:

$$c = \frac{4}{3}(x_1 - 1)^2 + \frac{1}{3}(x_2 + 1)^2 - \frac{1}{3}(x_1 - 1)(x_2 + 1)$$

Example


Contour Lines, Gaussian with Non-Diagonal Covariance
10.0
7.5
5.0
2.5
0.0
−2.5
−5.0
−7.5
−10.0
−10.0
−7.5
−5.0
−2.5
0.0
2.5
5.0
7.5
10.0