

Visualizing the Multivariate Normal, Lecture 9

Rebecca C. Steorts

September 15, 2015

Last class

- ▶ Class was done on the board – get notes if you missed lecture.
- ▶ Make sure to go through the Markdown example I posted on the webpage.

Spectral Decomposition

P is orthogonal if $P^T P = 1$ and $PP^T = 1$.

Theorem: Let A be symmetric $n \times n$. Then we can write

$$A = PDP^T,$$

where $D = \text{diag}(\lambda_1, \dots, \lambda_n)$ and P is orthogonal. The λ s are the eigenvalues of A and i th column of P is an eigenvector corresponding to λ_i .

Orthogonal matrices represent rotations of the coordinates.

Diagonal matrices represent stretchings/shrinkings of coordinates.

Properties

- ▶ The covariance matrix Σ is symmetric and positive definite, so we know from the spectral decomposition theorem that it can be written as

$$\Sigma = P\Lambda P^T.$$

- ▶ Λ is the diagonal matrix of the eigenvalues of Σ .
- ▶ P is the matrix whose columns are the orthonormal eigenvectors of Σ (hence V is an orthogonal matrix).
 - ▶ Geometrically, orthogonal matrices represent rotations.
 - ▶ Multiplying by P rotates the coordinate axes so that they are parallel to the eigenvectors of Σ .
 - ▶ Probabilistically, this tells us that the axes of the probability-contour ellipse are parallel to those eigenvectors.
 - ▶ The radii of those axes are proportional to the square roots of the eigenvalues.

Can we view the $\det(\Sigma)$ as a “variance”?

- ▶ Variance of one-dimensional Gaussian.
- ▶ From the SDT: $\det(\Sigma) = \prod_i \lambda_i$.
- ▶ Eigenvalues (λ_i) tell us how stretched or compressed the distribution is.
- ▶ View $\det(\Sigma)$ as stretching/compressing factor for the MVN density.
- ▶ We will see this from the contour plots later.

Our focus is visualizing MVN distributions in R.

What is a Contour Plot?

- ▶ Contour plot is a graphical technique for representing a 3-dimensional surface.
- ▶ We plot constant z slices (contours) on a 2-D format.
- ▶ The contour plot is an alternative to a 3-D surface plot.

The contour plot is formed by:

- ▶ Vertical axis: Independent variable 2.
- ▶ Horizontal axis: Independent variable 1.
- ▶ Lines: iso-response values.

Contour Plot

The lines of the contour plots denote places of equal probability mass for the MVN distribution

- ▶ The lines represent points of both variables that lead to the same height on the z-axis (the height of the surface)
- ▶ These contours can be constructed from the eigenvalues and eigenvectors of the covariance matrix
- ▶ The direction of the ellipse axes are in the direction of the eigenvalues
- ▶ The length of the ellipse axes are proportional to the constant times the eigenvector
- ▶ More specifically

$$||\Sigma^{-1/2}(\mathbf{X} - \boldsymbol{\mu})|| = c^2$$

has ellipsoids centered at $\boldsymbol{\mu}$ and axes at $\sqrt{(\lambda_i v_i)}$

Visualizing the MVN Distribution Using Contour Plots

The next figure below shows a contour plot of the joint pdf of a bivariate normal distribution. Note: we are plotting the theoretical contour plot. This particular distribution has mean

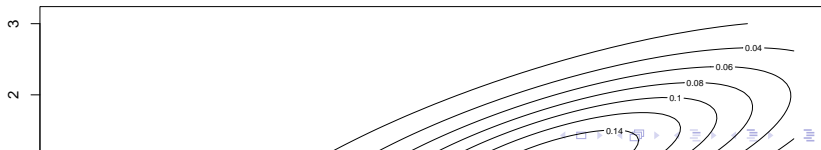
$$\boldsymbol{\mu} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

(solid dot), and variance matrix

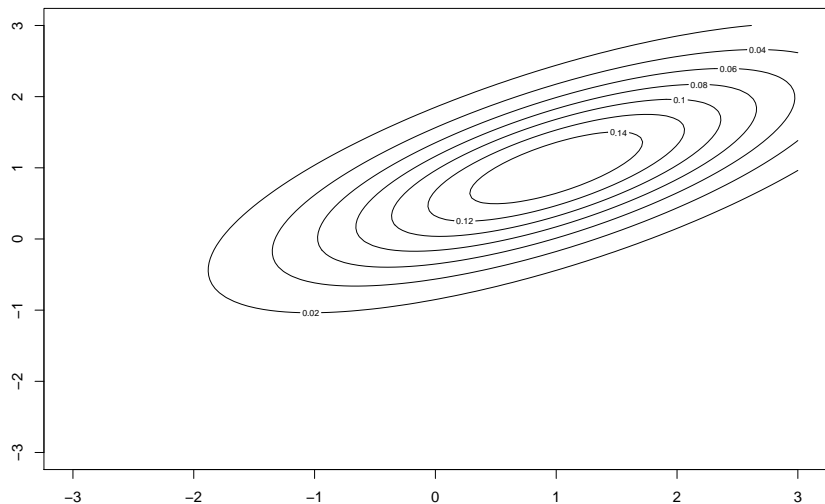
$$\boldsymbol{\Sigma} = \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix}.$$

Code to construct plot

```
library(mvtnorm)
x.points <- seq(-3,3,length.out=100)
y.points <- x.points
z <- matrix(0,nrow=100,ncol=100)
mu <- c(1,1)
sigma <- matrix(c(2,1,1,1),nrow=2)
for (i in 1:100) {
  for (j in 1:100) {
    z[i,j] <- dmvtnorm(c(x.points[i],y.points[j]),
                        mean=mu,sigma=sigma)
  }
}
contour(x.points,y.points,z)
```



Contour plot



Our findings

- ▶ Probability contours are ellipses.
- ▶ Density changes comparatively slowly along the major axis, and quickly along the minor axis.
- ▶ The two points marked $+$ in the figure have equal geometric distance from μ .
- ▶ But the one to its right lies on a higher probability contour than the one above it, because of the directions of their displacements from the mean.

Another example

Now consider the following samples drawn from bivariate normal densities:

$$\mathbf{x}_1, \dots, \mathbf{x}_n \stackrel{\text{iid}}{\sim} N_2 \left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 & 0.5 \\ 0.5 & 1 \end{pmatrix} \right)$$

and

$$\mathbf{y}_1, \dots, \mathbf{y}_n \stackrel{\text{iid}}{\sim} N_2 \left(\begin{pmatrix} -2 \\ -2 \end{pmatrix}, \begin{pmatrix} 1.5 & 1.5 \\ 1.5 & 1.5 \end{pmatrix} \right).$$

Kernel density estimation (KDE)

- ▶ KDE allows us to estimate the density from which each sample was drawn.
- ▶ This method (which you will learn about in other classes) allows us to approximate the density using a sample.
- ▶ There are R packages that use kde's such as `density()`.

Generate two bivariate normals

```
library(MASS)
bivn <- mvrnorm(1000, mu = c(0, 0),
               Sigma = matrix(c(1, .5, .5, 1), 2))
bivn2 <- mvrnorm(1000, mu = c(-2, 2),
                 Sigma = matrix(c(1.5, 1.5, 1.5, 1.5), 2))
```

Applying KDE and plotting

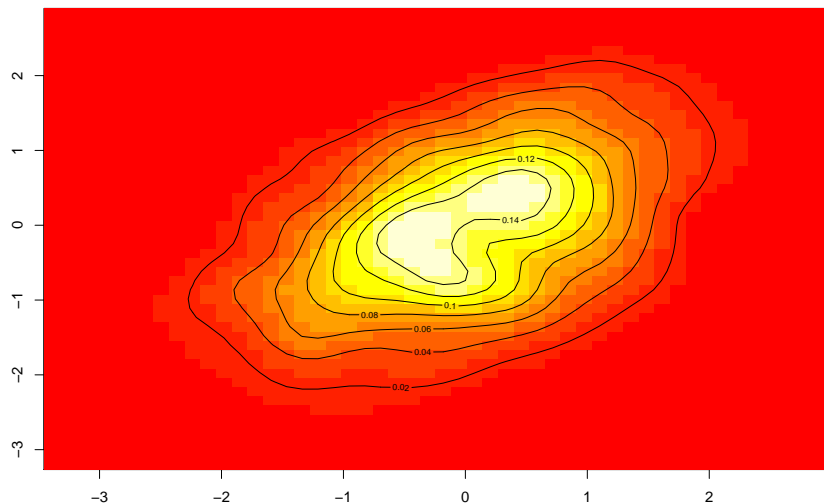
```
# now we do a kernel density estimate
bivn.kde <- kde2d(bivn[,1], bivn[,2], n = 50)
# now we do a mixture
bivn.mix <- kde2d(c(bivn[,1],bivn2[,1]),
                  c(bivn[,2],bivn2[,2]), n = 50)

##creating contour plots and saving them to files
pdf(file = "contour1.pdf")
image(bivn.kde); contour(bivn.kde, add = T)
dev.off()
```

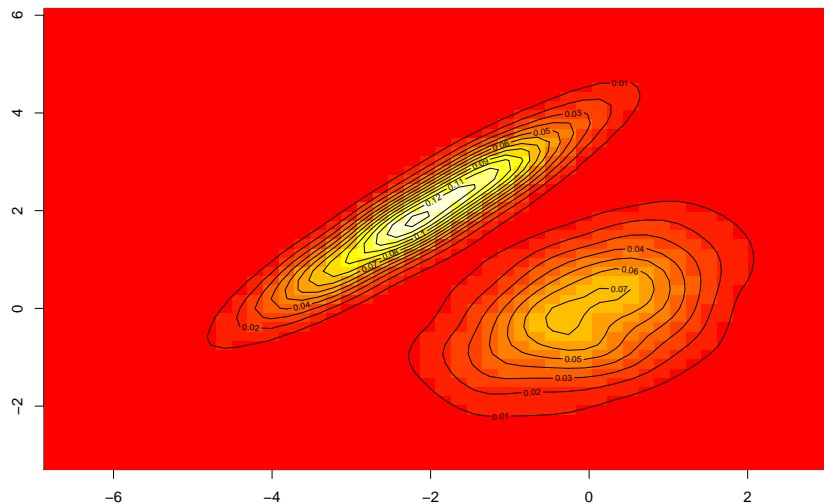
```
## pdf
## 2
```

```
pdf(file = "contour_both.pdf")
image(bivn.mix); contour(bivn.mix, add = T)
dev.off()
```


Contour plot of X



Contour plot of X and mixture of X, Y



What did we learn?

- ▶ The contour plot of \mathbf{X} (bivariate density): Color is the probability density at each point (red is low density and white is high density).
- ▶ Contour lines define regions of probability density (from high to low).
- ▶ Single point where the density is highest (in the white region) and the contours are approximately ellipses (which is what you expect from a Gaussian).

What can we say in general about the MVN density?

- ▶ The spectral decomposition theorem tells us that the contours of the multivariate normal distribution are ellipsoids.
- ▶ The axes of the ellipsoids correspond to eigenvectors of the covariance matrix.
- ▶ The radii of the ellipsoids are proportional to square roots of the eigenvalues of the covariance matrix.