The Multivariate Normal Distribution - Solutions

Exercise A - (5 min)

1. Load the [child test score dataset](https://ditraglia.com/data/child_test_data.csv) from our earlier lecture and use it to calculate the sample means, variance matrix, and correlation matrix of `mom.iq` and `kid.score`.
2. Continuing from part 2, create marginal and joint kernel density plots of `mom.iq` and `kid.score`. Do they appear to be normally distributed?
3. There's an R function called `cov2cor()` but there *isn't* one called `cor2cov()`. Why?
4. Reading from the color scale, the height of the "peak" in my two-dimensional kernel density plots from above was around 0.16. Why?
5. The contours of equal density for a pair of uncorrelated standard normal variables are circles. Why?

Solution

Parts 1-2

The variables `mom.iq` and `kid.score` do not appear to be normally distributed. Both are skewed and asymmetric.

```
library(tidyverse)
kids <- read_csv('https://ditraglia.com/data/child_test_data.csv')
dat <- kids |>
  select(mom.iq, kid.score)
colMeans(dat)
```

```
   mom.iq kid.score
100.00000  86.79724
```

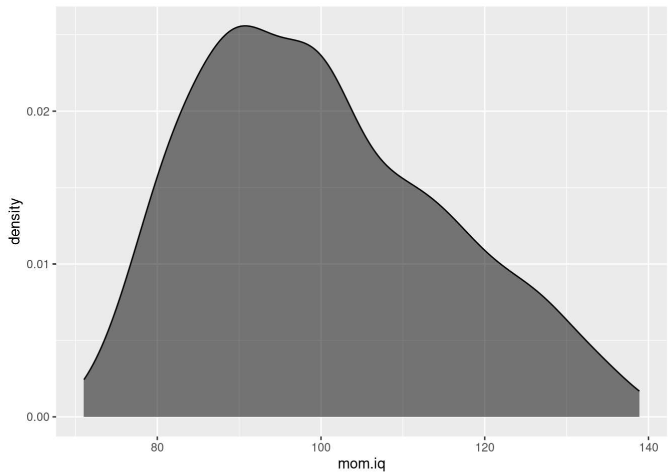
```
cov(dat)
```

```
             mom.iq kid.score
mom.iq    225.0000  137.2443
kid.score 137.2443  416.5962
```

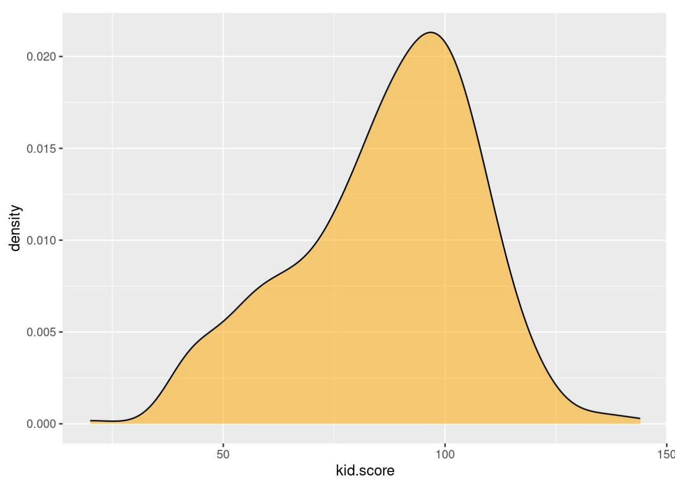
```
cor(dat)
```

```
             mom.iq kid.score
mom.iq    1.0000000 0.4482758
kid.score 0.4482758 1.0000000
```

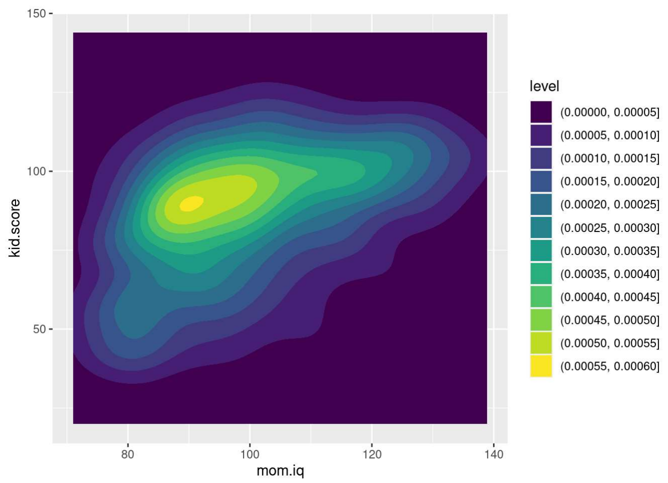
```
ggplot(kids) +
  geom_density(aes(x = mom.iq), fill = 'black', alpha = 0.5)
```




```
ggplot(kids) +
  geom_density(aes(x = kid.score), fill = 'orange', alpha = 0.5)
```




```
ggplot(kids) +
  geom_density2d_filled(aes(x = mom.iq, y = kid.score))
```



Part 3

A correlation matrix contains *strictly less information* than a covariance matrix. If I give you the correlation matrix of (X_1, X_2), then I haven't told you the variances of either X_1 or X_2. This means you can't go from a correlation matrix to a covariance matrix, although you can go in the reverse direction.

Part 4

If Z_1 and Z_2 are independent standard normal random variables, then their joint density equals the product of their marginal densities:

$$\begin{aligned} f(z_1, z_2) = f(z_1)f(z_2) &= \frac{1}{\sqrt{2\pi}}\exp\left(-\frac{z_1^2}{2}\right) \times \frac{1}{\sqrt{2\pi}}\exp\left(-\frac{z_2^2}{2}\right) \\ &= \frac{1}{2\pi}\exp\left\{-\frac{1}{2}(z_1^2 + z_2^2)\right\}. \end{aligned}$$

Since $1/(2\pi)$ is positive and $-1/2$ is negative, this function is maximized when $z_1 + z_2^2$ is made as small as possible, i.e. at $(0, 0)$. Substituting these gives $f(0, 0) = 1/(2\pi) \approx 0.159$.

Part 5

From the expression in the previous solution, $f(z_1, z_2)$ is constant whenever $(z_1^2 + z_2^2)$ is constant, and the expression $(z_1^2 + z_2^2) = C$ describes a circle centered at $(0, 0)$ for any positive constant C.

Exercise B - (5 min)

Suppose that $Z_1, Z_2 \sim$ iid $\text{N}(0,1)$ and

$$\boldsymbol{X} = \mathbf{A}\boldsymbol{Z}, \quad \boldsymbol{X} = \begin{bmatrix} X_1 \\ X_2 \end{bmatrix}, \quad \mathbf{A} = \begin{bmatrix} a & b \\ c & d \end{bmatrix}, \quad \begin{bmatrix} Z_1 \\ Z_2 \end{bmatrix}.$$

Calculate $\text{Var}(X_1)$, $\text{Var}(X_2)$, and $\text{Cov}(X_1, X_2)$ in terms of the constants a, b, c, d. Using these calculations, show that the variance-covariance matrix of $\boldsymbol{X}$ equals $\mathbf{A}\mathbf{A}'$. Use this result to work out the variance-covariance matrix of my example from above with $a = 2, b = 1, c = 1, d = 4$ and check that it agrees with the simulations.

Solution

First we'll calculate the variances. Since Z_1 and Z_2 are independent and both have a variance of one,

$$\text{Var}(X_1) = \text{Var}(aZ_1 + bZ_2) = a^2\text{Var}(Z_1) + b^2\text{Var}(Z_2) = a^2 + b^2$$

Analogously, $\text{Var}(X_2) = c^2 + d^2$. Next we'll calculate the covariance. Again, since Z_1 and Z_2 are independent,

$$\begin{aligned} \text{Cov}(X_1, X_2) &= \text{Cov}(aZ_1 + bZ_2, cZ_1 + dZ_2) \\ &= \text{Cov}(aZ_1, cZ_1) + \text{Cov}(bZ_2, dZ_2) \\ &= ac\,\text{Var}(Z_1) + bd\,\text{Var}(Z_2) \\ &= ac + bd \end{aligned}$$

Now we collect these results into a matrix, the variance-covariance matrix of $\mathbf{X}$:

$$\text{Var}(\mathbf{X}) \equiv \begin{bmatrix} \text{Var}(X_1) & \text{Cov}(X_1, X_2) \\ \text{Cov}(X_1, X_2) & \text{Var}(X_2) \end{bmatrix} = \begin{bmatrix} a^2 + b^2 & ac + bd \\ ac + bd & c^2 + d^2 \end{bmatrix}$$

Multiplying through, this is precisely equal to $\mathbf{A}\mathbf{A}'$

$$\mathbf{A}\mathbf{A}' = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} a & c \\ b & d \end{bmatrix} = \begin{bmatrix} a^2 + b^2 & ac + bd \\ ac + bd & c^2 + d^2 \end{bmatrix}$$

Finally, substituting the values from the example above

```
A <- matrix(c(2, 1,
              1, 4), byrow = TRUE, ncol = 2, nrow = 2)
A %*% t(A)
```

```
     [,1] [,2]
[1,]    5    6
[2,]    6   17
```

The sample variance-covariance matrix from our simulations is quite close:

```
set.seed(99999)
n <- 1e5
z1 <- rnorm(n)
z2 <- rnorm(n)
z <- cbind(z1, z2)
rm(z1, z2)
x <- cbind(x1 = 2 * z[, 1] +     z[, 2],
           x2 =     z[, 1] + 4 * z[, 2])
var(x)
```

```
         x1        x2
x1 5.031082  6.031717
x2 6.031717 17.068951
```

Exercise C - (15 min)

1. Suppose we wanted the correlation between X_1 and X_2 to be ρ, a value that might not equal 0.5. Modify the argument from above accordingly.
2. In our discussion above, X_1 and X_2 both had variance equal to one, so their correlation equaled their covariance. More generally, we may want to generate X_1 with variance σ_1^2 and X_2 with variance σ_2^2, where the covariance between them equals σ_{12}. Extend your reasoning from the preceding exercise to find an appropriate matrix $\mathbf{A}$ such that $\mathbf{A}\mathbf{Z}$ has the desired variance-covariance matrix.
3. Check your calculations in the preceding part by transforming the simulations `z` from above into `x` such that `x1` has variance one, `x2` has variance four, and the correlation between them equals 0.4. Make a density plot of your result.

Solution

Part 1

All we have to do is replace 0.5 with ρ. Everything else goes through as before:

$$\begin{bmatrix} 1 & 0 \\ \rho & \sqrt{1-\rho^2} \end{bmatrix} \begin{bmatrix} Z_1 \\ Z_1 \end{bmatrix} \sim \text{N}\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 & \rho \\ \rho & 1 \end{bmatrix} \right).$$

Part 2

We just have to work out the appropriate values of the constants a, b, c, and d in the equations $X_1 = aZ_1 + bZ_2$ and $X_2 = cZ_1 + dZ_2$. Because there are *many* matrices $\mathbf{A}$ that will do the trick, we'll adopt the convention that $b = 0$ and $a, d > 0$, as above.

Since $b = 0$, we have $X_1 = aZ_1$. Since Z_1 is a standard normal, to give X_1 a variance of σ_1^2 we need to set $a = \sigma_1$. By our convention, this is the *positive* square root of σ_1^2 so that $a > 0$. As we calculated in an earlier exercise, $\text{Cov}(X_1, X_2) = ac + bd$. Since $a = \sigma_1$ and $b = 0$, this simplifies to $\text{Cov}(X_1, X_2) = \sigma_1 c$. In order for this covariance to equal σ_{12}, we need to set $c = \sigma_{12}/\sigma_1$.

All that remains is to set the variance of X_2 to σ_2^2. Again using a calculation from a previous exercise, $\text{Var}(X_2) = c^2 + d^2$. Substituting our solution for c and equating to σ_2^2 gives $d^2 = \sigma_2^2 - \sigma_{12}^2/\sigma_1^2$. To ensure that $d > 0$ we set it equal to the positive square root: $d = \sqrt{\sigma_2^2 - \sigma_{12}^2/\sigma_1^2}$. In matrix form:

$$\begin{bmatrix} \sigma_1 & 0 \\ \sigma_{12}/\sigma_1 & \sqrt{\sigma_2^2 - \sigma_{12}^2/\sigma_1^2} \end{bmatrix} \begin{bmatrix} Z_1 \\ Z_1 \end{bmatrix} \sim \text{N}\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \sigma_1^2 & \sigma_{12} \\ \sigma_{12} & \sigma_2^2 \end{bmatrix} \right).$$

Part 3

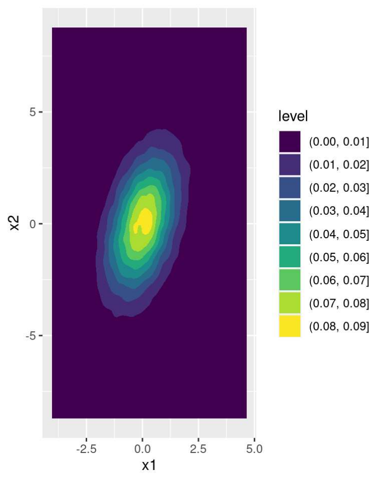
```
s1 <- 1
s2 <- 2
r <- 0.4
s12 <- r * (s1 * s2)
A <- matrix(c(s1, 0,
              s12 / s1, sqrt(s2^2 - s12^2 / s1^2)),
            byrow = TRUE, nrow = 2, ncol = 2)
x <- t(A %*% t(z))
colnames(x) <- c('x1', 'x2')
cov(x)
```

```
          x1        x2
x1 1.0063783 0.8059712
x2 0.8059712 4.0178160
```

```
cor(x)
```

```
          x1        x2
x1 1.0000000 0.4008149
x2 0.4008149 1.0000000
```

```
as_tibble(x) |>
  ggplot(aes(x1, x2)) +
  geom_density2d_filled() +
  coord_fixed()
```



Exercise D - (∞ min)

1. To check if a (3×3) matrix `M` is p.d., proceed as follows. First check that `M[1,1]` is positive. Next use `det()` to check that the *determinant* of `M[1:2,1:2]` is positive. Finally check that `det(M)` is positive. If `M` passes all the tests, it's p.d. The same procedure works for any matrix: check that the determinant of each leading principal minor is positive. Only one of these matrices is p.d. Which one?

$$\mathbf{A} = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 2 & 1 \\ 3 & 1 & 3 \end{bmatrix}, \quad \mathbf{B} = \begin{bmatrix} 3 & 2 & 1 \\ 2 & 3 & 1 \\ 1 & 1 & 3 \end{bmatrix}$$

2. Let $\mathbf{\Sigma}$ be the p.d. matrix from part 1. Use `chol()` to make 100,000 draws from a MV normal distribution with this variance matrix. Check your work with `var()`.
3. Install the package `mvtnorm()` and consult `?rmvnorm()`. Then repeat the preceding exercise "the easy way," without using `chol()`. Check your work.
4. Let $Y = \alpha X_1 + \beta X_2$, $\mathbf{v} = (\alpha, \beta)'$, and $\mathbf{\Sigma} = \text{Var}(X_1, X_2)$. Show that $\text{Var}(Y) = \mathbf{v}'\mathbf{\Sigma}\mathbf{v}$.

Solution

Part 1

Both $\mathbf{A}$ and $\mathbf{B}$ are symmetric. The matrix $\mathbf{A}$ is *not* positive definite because its determinant is negative:

```r
A <- matrix(c(1, 2, 3,
              2, 2, 1,
              3, 1, 3),
            byrow = TRUE, nrow = 3, ncol = 3)
det(A)
```

```
[1] -13
```

The matrix $\mathbf{B}$ is positive definite since `B[1, 1]` is positive, the determinant of `B[1:2, 1:2]` is positive, and the determinant of `B` itself is positive:

```r
B <- matrix(c(3, 2, 1,
              2, 3, 1,
              1, 1, 3),
            byrow = TRUE, nrow = 3, ncol = 3)
det(B[1:2, 1:2])
```

```
[1] 5
```

```r
det(B)
```

```
[1] 13
```

Part 2

```r
R <- chol(B)
L <- t(R)
n_sims <- 1e5
set.seed(29837)
z <- matrix(rnorm(3 * n_sims), nrow = n_sims, ncol = 3)
x <- t(L %*% t(z))
cov(x)
```

```
         [,1]     [,2]     [,3]
[1,] 2.986340 1.985103 1.006587
[2,] 1.985103 2.983812 1.001514
[3,] 1.006587 1.001514 2.997953
```

Part 3

```r
#install.packages('mvtnorm')
library(mvtnorm)
set.seed(29837)
x_alt <- rmvnorm(n_sims, sigma = B)
cov(x_alt)
```

```
         [,1]     [,2]     [,3]
[1,] 2.984944 1.997268 1.003196
[2,] 1.997268 2.992995 1.006207
[3,] 1.003196 1.006207 3.010338
```

Part 4

$$
\begin{aligned}
\mathbf{v}'\boldsymbol{\Sigma}\mathbf{v} &= \begin{bmatrix}\alpha & \beta\end{bmatrix}\begin{bmatrix}\sigma_1^2 & \sigma_{12}\\ \sigma_{12} & \sigma_2^2\end{bmatrix}\begin{bmatrix}\alpha\\ \beta\end{bmatrix}\\
&= \alpha^2\mathrm{Var}(X_1) + 2\alpha\beta\,\mathrm{Cov}(X_1, X_2) + \beta^2\mathrm{Var}(X_2)\\
&= \mathrm{Var}(\alpha X_1 + \beta X_2) = \mathrm{Var}(Y).
\end{aligned}
$$