

26.8: Appendix

26.8.1 Estimating linear regression parameters

We generally estimate the parameters of a linear model from data using *linear algebra*, which is the form of algebra that is applied to vectors and matrices. If you aren't familiar with linear algebra, don't worry – you won't actually need to use it here, as R will do all the work for us. However, a brief excursion in linear algebra can provide some insight into how the model parameters are estimated in practice.

First, let's introduce the idea of vectors and matrices; you've already encountered them in the context of R, but we will review them here. A matrix is a set of numbers that are arranged in a square or rectangle, such that there are one or more *dimensions* across which the matrix varies. It is customary to place different observation units (such as people) in the rows, and different variables in the columns. Let's take our study time data from above. We could arrange these numbers in a matrix, which would have eight rows (one for each student) and two columns (one for study time, and one for grade). If you are thinking “that sounds like a data frame in R” you are exactly right! In fact, a data frame is a specialized version of a matrix, and we can convert a data frame to a matrix using the `as.matrix()` function.

```
df <-  
  tibble(  
    studyTime = c(2, 3, 5, 6, 6, 8, 10, 12) / 3,  
    priorClass = c(0, 1, 1, 0, 1, 0, 1, 0)  
  ) %>%  
  mutate(  
    grade =  
      studyTime * betas[1] +  
      priorClass * betas[2] +  
      round(rnorm(8, mean = 70, sd = 5))  
  )  
  
df_matrix <-  
  df %>%  
  dplyr::select(studyTime, grade) %>%  
  as.matrix()
```

We can write the general linear model in linear algebra as follows:

$$Y = X * \beta + E$$

This looks very much like the earlier equation that we used, except that the letters are all capitalized, which is meant to express the fact that they are vectors.

We know that the grade data go into the Y matrix, but what goes into the

Unexpected text node: '26.7).'

Independent variable		Design matrix	
0.67		77	1
1.0		85	1
1.67		84	1
2.0		80	1
2.0	=	90	1
2.67		77	1
3.33		98	1
4.0		93	1

$$* \beta + E$$

Figure 26.7: A depiction of the linear model for the study time data in terms of matrix algebra.

The rules of matrix multiplication tell us that the dimensions of the matrices have to match with one another; in this case, the design matrix has dimensions of 8 (rows) X 2 (columns) and the Y variable has dimensions of 8 X 1. Therefore, the β matrix needs to have dimensions 2 X 1, since an 8 X 2 matrix multiplied by a 2 X 1 matrix results in an 8 X 1 matrix (as the matching middle dimensions drop out). The interpretation of the two values in the β matrix is that they are the values to be multiplied by study time and 1 respectively to obtain the estimated grade for each individual. We can also view the linear model as a set of individual equations for each individual:

$$\hat{y}_1 = studyTime_1 * \beta_1 + 1 * \beta_2$$

$$\hat{y}_2 = studyTime_2 * \beta_1 + 1 * \beta_2$$

...

$$\hat{y}_8 = studyTime_8 * \beta_1 + 1 * \beta_2$$

Remember that our goal is to determine the best fitting values of β given the known values of X and Y . A naive way to do this would be to solve for β using simple algebra – here we drop the error term E because it's out of our control:

$$\hat{\beta} = \frac{Y}{X}$$

The challenge here is that X and β are now matrices, not single numbers – but the rules of linear algebra tell us how to divide by a matrix, which is the same as multiplying by the *inverse* of the matrix (referred to as X^{-1}). We can do this in R:

```
# compute beta estimates using linear algebra

#create Y variable 8 x 1 matrix
Y <- as.matrix(df$grade)
#create X variable 8 x 2 matrix
X <- matrix(0, nrow = 8, ncol = 2)
#assign studyTime values to first column in X matrix
X[, 1] <- as.matrix(df$studyTime)
#assign constant of 1 to second column in X matrix
X[, 2] <- 1

# compute inverse of X using ginv()
# %*% is the R matrix multiplication operator

beta_hat <- ginv(X) %*% Y #multiple the inverse of X by Y
print(beta_hat)
```

```
##      [,1]
## [1,]  4.3
## [2,] 76.0
```

Anyone who is interested in serious use of statistical methods is highly encouraged to invest some time in learning linear algebra, as it provides the basis for nearly all of the tools that are used in standard statistics.

This page titled [26.8: Appendix](#) is shared under a [CC BY-NC 4.0](#) license and was authored, remixed, and/or curated by [Russell A. Poldrack](#) via [source content](#) that was edited to the style and standards of the LibreTexts platform.