

# Projections Onto Subspaces

Jeannie Ramelb

May 10, 2017

## 1 Introduction

When I spoke to Professor Demeo, he recommend that I do my medium presentation on a topic that was more challenging. I made it pretty obvious in my first presentation that I did not want to pursue a career in mathematics. I told him about the work I do in the insurance industry. He suggested that I talk about projections onto subspaces and ordinary least squares (linear regression) because it can be applied in the insurance industry. I took the challenge and that is why I did my medium 20-minute presentation on projections onto subspaces.

## 2 Projections

We can represent systems of linear equations in terms of matrices and vectors.

Let  $L = \text{span}(\vec{v})$

$$L = \{a * \vec{v} | a \in R\}$$

where a is all possible scalar multiples.

$\text{proj}_L(\vec{w})$  is the shadow of  $\vec{w}$  onto L. Since  $\vec{v}$  is the spanning vector of the line L then the projection can be expressed as

$$\text{proj}_L(\vec{w}) = \frac{\vec{w} \cdot \vec{v}}{\vec{v} \cdot \vec{v}} \cdot \vec{v}$$

### 3 Projections on Subspaces

If  $V$  is a subspace of  $R^n$  then  $V^\perp$  also subspace.  $\vec{w} \in R^n$  then  $\vec{w} = \vec{v} + \vec{x}$  where  $\vec{v} \in V$  and  $\vec{x} \in V^\perp$ .

We can define  $proj_V(\vec{w}) = \vec{v}$  which is part of  $\vec{w}$  that includes  $\vec{v} \in V$ . We can define  $proj_{V^\perp}(\vec{w}) = \vec{x}$  which is part of  $\vec{w}$  that includes the orthogonal complement  $\vec{x} \in V^\perp$ . This is also a linear transformation.  $V$  is a subspace of  $R^n$ .  $\{\vec{s}_1, \vec{s}_2, \dots, \vec{s}_k\}$  is a basis for  $V$  set. Then if you take the span of vectors you can construct any vectors in that subspace and those vectors are linearly independent.

The projection of a vector onto subspace is the closest point to another vector contained in  $V$ .

### 4 Ordinary Least Squares Regression-Linear Model

We collect data and observation. Linear regression takes this information and summarizes it by drawing a line that represents the data.

At the origin, the formula of OLS is

$$y_n = \sum_{i=0}^k m_i x_{ni} + \epsilon_n \quad (1)$$

The  $x$  variable is considered to be an explanatory variable where as  $y$  variable is considered to be a dependent variable.  $m$  is found by minimizing error or prediction.

For example, if we take the  $x$ -axis to be traveling and  $y$ -axis to be happiness, the formula would be  $y = b + mx + \epsilon$ , where  $y$  is happiness,  $b$  is  $y$ -intercept, and  $\epsilon$  is the error.  $b$  tells you what value of happiness would be if the value of independent variable was 0.  $m$  says by how much the value of happiness changes in response to a change in the value of traveling.

Linear regression takes all of the values, squares them, and sums them all up

$$\sum_{i=1}^n (Y_i - \hat{Y}_i)^2$$

or

$$\sum_{i=1}^n (Y_i - (mx_i + b))^2$$

This will spit out a positive value because of the positive distance. We get this line that best represents the data "by minimizing the  $y$ -distances from the corresponding  $y$ -values on the to-be-determined line. The square of each of those distances is of the form  $(y_j - (mx_j + b))^2$ ." The total distance is considered the square root of the sum of these squares.

To relate this back to projections, linear regression is exactly the projection of  $y$  values onto subspace of the span of two vectors  $\vec{v}_1$  and  $\vec{v}_2$ , where  $\vec{v}_1$  is the vector in  $\mathbb{R}^2$  with all coordinates in 1 and  $\vec{v}_2$  is the vector in  $\mathbb{R}^2$  with coordinates  $x_j$ . Minimizing the square of distance from  $y$  to the vector in subspace of the span of vectors is the same as finding the orthogonal projection.

## 5 Ordinary Least Squares Regression-Multiple

I provided a slide with a picture on this, however, I did not elaborate on it. This OLS method can be extended to models that involve 2 or more explanatory variables. The algebra becomes really complex because of the multiple explanatory variables.

## 6 Linear Regression Applications

Linear regression helps us find out where a business is going and growing and is essential for achieving long term goals. Linear Regression can be used in business to evaluate trends and make predictions. We use the example of sales. Trend line using linear regression can be created and the company may use the slope to estimate the future sales.

Linear regression is used in economics to determine how consumers act when there has been a change in price. Linear regression helps guide decisions on pricing an item based on demand.

Linear regression is particularly used in implication for research. Researchers can use small prediction models to talk about a larger population sample. If what we can predict of what a smaller population does, we can predict what a larger population does. This can save an institution or a researcher a lot of money.

Linear regression can be used to analyze a risk. For example in auto insurance, linear regression pulls together a homogeneous units which are similar types of risk into a pool, charges each person in the pool a small premium in exchange for protection. An underwriter uses this method to determine how likely it is that a loss will occur. With linear regression, actuaries can actually predict how likely a loss is to occur!

## 7 Works Cited

Cherlin, Gregory. "Lecture 25 Orthogonal Projection and Least Squares Approximation." Cherlin Math 250. Rutgers, The State University of New Jersey, n.d. Web. 11 May 2017.

Linear Regression. Yale University, n.d. Web. 11 May 2017.

Shitut, Neha. "Popular Applications of Linear Regression for Businesses." Analytics Training. Jigsaw Academy Education Pvt Ltd, 11 Apr. 2017. Web. 11 May 2017.