

Linear Algebra and Matrix Methods - Part III: Lecture V

Charles B. Moss ¹

¹University of Florida

January 18, 2012

- 1 Readings
- 2 Vector Spaces
 - Linear Independence and Dependence
- 3 Orthonormal Bases and Projections
- 4 Projection Matrices
- 5 Readings for Next Lecture

Readings

- Anderson, T.W. 1984. *An Introduction to Multivariate Statistical Analysis* Second Edition. John Wiley & Sons. (A.5 pp. 603-608).
- * Dhrymes, P.J. 2000. *Mathematics for Econometrics* Third Edition. Springer-Verlag. (Chapter 2 [Section 3 - 4]).
- * Greene, W.H. 2012. *Econometric Analysis* Seventh Edition. Prentice Hall. (Appendix A: Matrix Algebra pp. 973-1014).
- * Popper, K. 2010. *The Logic of Scientific Discovery* Routledge Classics [Chapter 5: The Problem of the Empirical Basis, 74-94].

Vector Spaces

- Definitions

- **Definition 2.1.** Let S be a collection of $m \times 1$ vectors satisfying the following
 - $x_1 \in S$ and $x_2 \in S$, then $x_1 + x_2 \in S$.
 - if $x \in S$ and α is a real scalar then $\alpha x \in S$.

Then S is called a vector space in m -dimensional space. If S is a subset of T , which is another vectors space in m -dimensional space, S is called a vector subspace of T .

Vector Spaces - Continued

- **Definition 2.2** Let $\{x_1, \dots, x_n\}$ be a set of $m \times 1$ vectors in the vector space S . If each vector in S can be expressed as a linear combination of the vectors $x_1, \dots, x_n$, then the set $\{x_1, \dots, x_n\}$ is said to span or generate the vector space S , and $\{x_1, \dots, x_n\}$ is called a spanning set of S .

Linear Independence and Dependence

- **Definition 2.6** The set of $m \times 1$ vectors $\{x_1, \dots, x_n\}$ is said to be a linearly independent if the only solution to the equation

$$\sum_{i=1}^n \alpha_i x_i = 0 \quad (1)$$

is the zero vector $\alpha_1 = \dots = \alpha_n = 0$.

Linear Independence and Dependence

- Going back to the matrix

$$\begin{pmatrix} 1 & 9 & 5 \\ 3 & 8 & 8 \\ 4 & 16 & 13 \end{pmatrix} \Rightarrow \begin{pmatrix} 1 & 0 & 37/20 \\ 0 & 1 & 7/20 \\ 0 & 0 & 0 \end{pmatrix} \quad (2)$$

This reduction implies that

$$\frac{37}{20} \begin{pmatrix} 1 \\ 3 \\ 4 \end{pmatrix} + \frac{7}{20} \begin{pmatrix} 9 \\ 7 \\ 16 \end{pmatrix} = \begin{pmatrix} 5 \\ 8 \\ 13 \end{pmatrix} \quad (3)$$

Or that the third column of the matrix is a linear combination of the first two.

Orthonormal Basis and Projects

- Suppose that a set of vectors $\{x_1, \dots, x_r\}$ for a basis for some space S in R^m space such that $r \leq m$. For mathematical simplicity, we may want to form an orthogonal basis for this space. One way to form such a basis is the Gram-Schmidt orthonormalization. In this procedure, we want to generate a new set of vectors $\{y_1, \dots, y_r\}$ that are orthonormal.
- The Gram-Schmidt process is

$$\begin{aligned}y_1 &= x_1 \\y_2 &= x_2 - \frac{x_2' y_1}{y_1' y_1} y_1 \\y_3 &= x_3 - \frac{x_3' y_1}{y_1' y_1} y_1 - \frac{x_3' y_2}{y_2' y_2} y_2\end{aligned} \tag{4}$$

Orthonormal Basis and Projects

- Which produces a set of orthogonal vectors, and then

$$z_i = \frac{y_i}{\sqrt{y_i' y_i}} \quad (5)$$

Orthonormal Basis and Projects - Continued

- Example, the vectors

$$x_1 = \begin{pmatrix} 1 \\ 3 \\ 4 \end{pmatrix}, x_2 = \begin{pmatrix} 9 \\ 7 \\ 16 \end{pmatrix} \quad (6)$$

span a plane in three dimension space. Setting $y_1 = x_1$, y_2 is derived as

$$y_2 = \begin{pmatrix} 9 \\ 7 \\ 16 \end{pmatrix} - \frac{\begin{pmatrix} 9 & 7 & 16 \end{pmatrix} \begin{pmatrix} 1 \\ 3 \\ 4 \end{pmatrix}}{\begin{pmatrix} 1 & 3 & 4 \end{pmatrix} \begin{pmatrix} 1 \\ 3 \\ 4 \end{pmatrix}} \begin{pmatrix} 1 \\ 3 \\ 4 \end{pmatrix} = \begin{pmatrix} 70/13 \\ -50/13 \\ 20/13 \end{pmatrix} \quad (7)$$

Orthonormal Basis and Projects - Continued

- The vectors can then be normalized to one. However, to test for orthogonality

$$\begin{pmatrix} 1 & 3 & 4 \end{pmatrix} \begin{pmatrix} 70/13 \\ -50/13 \\ 20/13 \end{pmatrix} = 0 \quad (8)$$

- Theorem 2.13** Every r -dimensional vector space, except the zero-dimensional space $\{0\}$, has an orthonormal basis.

Orthonormal Basis and Projects - Continued

- **Theorem 2.14** Let $\{z_1, \dots, z_r\}$ be an orthonormal basis for some vector space S , of R^m . Then each $x \in R^m$ can be expressed uniquely as

$$x = u + v \tag{9}$$

where $u \in S$ and v is a vector that is orthogonal to every vector in S .

Orthonormal Basis and Projects - Continued

- **Definition 2.10** Let S be a vector subspace of R^m . The orthogonal complement of S , denoted $S^\perp$, is the collection of all vectors in R^m that are orthogonal to every vector in S : That is, $S^\perp = \{x : x \in R^m \text{ and } x'y = 0, \forall y \in S\}$.
- **Theorem 2.15** If S is a vector subspace of R^m then its orthogonal complement $S^\perp$ is also a vector subspace of R^m .

Projection Matrices

- The orthogonal projection of an $m \times 1$ vector x onto a vector space S can be expressed in matrix form.
 - Let $\{z_1, \dots, z_r\}$ be any orthonormal basis for S while $\{z_1, \dots, z_m\}$ is an orthonormal basis for R^m . Any vector x can be written as

$$x = (\alpha_1 z_1 + \dots + \alpha_r z_r) + (\alpha_{r+1} z_{r+1} + \dots + \alpha_m z_m) = u + v \quad (10)$$

Projection Matrices

- Aggregating $\alpha = \begin{pmatrix} \alpha'_1 & \alpha'_2 \end{pmatrix}'$ where $\alpha_1 = \begin{pmatrix} \alpha_1 & \cdots & \alpha_r \end{pmatrix}$ and $\alpha_2 = \begin{pmatrix} \alpha_{r+1} & \cdots & \alpha_m \end{pmatrix}$ and assuming a similar decomposition of $Z = \begin{bmatrix} Z_1 & Z_2 \end{bmatrix}$, the vector x can be written as

$$X = Z\alpha = Z_1\alpha_1 + Z_2\alpha_2$$

$$u = Z_1\alpha_1 \tag{11}$$

$$v = Z_2\alpha_2$$

Projection Matrices - Continued

- Given orthogonality, we know that $Z_1'Z_1 = I_r$ and $Z_1'Z_2 = [0]$, and so

$$Z_1Z_1'x = Z_1Z_1' \begin{bmatrix} Z_1 & Z_2 \end{bmatrix} \begin{bmatrix} \alpha_1 \\ \alpha_2 \end{bmatrix} = \begin{bmatrix} Z_1 & 0 \end{bmatrix} \begin{bmatrix} \alpha_1 \\ \alpha_2 \end{bmatrix} = Z_1\alpha_1 = u \quad (12)$$

- Theorem 2.17** Suppose the columns of the $m \times r$ matrix Z_1 form an orthonormal basis for the vector space S which is a subspace of R^m . If $x \in R^m$, the orthogonal projection of x onto S is given by $Z_1Z_1'x$.
- Projection matrices allow the division of the space into a spanned space and a set of orthogonal deviations from the spanning set. One such separation involves the Gram-Schmidt system.

Projection Matrices - Continued

- In general, if we define the $m \times r$ matrix $X_1 = (x_1, \dots, x_r)$ and define the linear transformation of this matrix that produces an orthonormal basis as A , so that

$$Z_1 = X_1 A \quad (13)$$

we are left with the result that

$$Z_1' Z_1 = A' X_1' X_1 A = I_r \quad (14)$$

Given that the matrix A is nonsingular, the projection matrix that maps any vector x onto the spanning set then becomes

$$P_S = Z_1 Z_1' = X_1 A A' X_1' = X_1 (X_1' X_1)^{-1} X_1'. \quad (15)$$

Projection Matrices - Continued

- Ordinary least squares is also a spanning decomposition. In the traditional linear model

$$y = X\beta + \epsilon \quad (16)$$

$$\hat{y} = X\hat{\beta}$$

within this formulation β is chosen to minimize the error between y and estimated $\hat{y}$

$$(y - X\beta)'(y - X\beta). \quad (17)$$

Projection Matrices - Continued

- This problem implies minimizing the distance between the observed y and the predicted plane $X\beta$, which implies orthogonality. If X has full column rank, the projection space becomes $X(X'X)^{-1}X'$ and the projection then becomes

$$X\beta = X(X'X)^{-1}X'y \quad (18)$$

Premultiplying each side by X' yields

$$\begin{aligned} X'X\beta &= X'X(X'X)^{-1}X'y \\ \beta &= (X'X)^{-1}X'X(X'X)^{-1}X'y \end{aligned} \quad (19)$$

$$\beta = (X'X)^{-1}X'y$$

Projection Matrices - Continued

- Idempotent matrices can be defined as any matrix such that $AA = A$.
 - Note that the sum of square errors under the projection can be expressed as

$$\begin{aligned}SSE &= (y - X\beta)' (y - X\beta) \\&= \left(y - X (X'X)^{-1} X'y \right)' \left(y - X (X'X)^{-1} X'y \right) \\&= \left(\left(I_n - X (X'X)^{-1} X' \right) y \right)' \left(I_n - X (X'X)^{-1} X' \right) y \\&= y' \left(I_n - X (X'X)^{-1} X' \right) \left(I_n - X (X'X)^{-1} X' \right) y\end{aligned}\tag{20}$$

Projection Matrices - Continued

- Continued

- In general, the matrix $I_n - X(X'X)^{-1}X'$ is referred to as an idempotent matrix.

$$\begin{aligned} & \left(I_n - X(X'X)^{-1}X' \right) \left(I_n - X(X'X)^{-1}X' \right) = \\ & I_n - X(X'X)^{-1}X' - X(X'X)^{-1}X' + X(X'X)^{-1}X'X(X'X)^{-1}X' \quad (21) \\ & = I_n - X(X'X)^{-1}X' \end{aligned}$$

Projection Matrices - Continued

- Continued
 - Thus, the SSE can be expressed as

$$\begin{aligned}SSE &= y' \left(I_n - X (X'X)^{-1} X' \right) y \\ &= y'y - y'X (X'X)^{-1} X'y = v'v\end{aligned}\tag{22}$$

which is the sum of the orthogonal errors from the regression.

Readings for Next Lecture

- * Khuri, A.I. 1993. *Advanced Calculus with Applications in Statistics*. John Wiley & Sons. (Chapters 3, 4, 7, and 8)
- * Moss, C.B. 2010. *Static and Dynamic Optimization for Agricultural and Resource Economics* Draft Textbook. (Chapter 2 pp. 15-40)
- * Popper, K. 2010. *The Logic of Scientific Discovery* Routledge Classics [Chapter 6: Degrees of Testability, 95-120].