

Notes for ECE269 - Linear Algebra

Chapter 6

Sam Cowin

July 27, 2020

- 1 Linear Equations in Linear Algebra**
- 2 Matrix Algebra**
- 3 Determinants**
- 4 Vector Spaces**
- 5 Eigenvalues and Eigenvectors**
- 6 Orthogonality and Least Squares**

There are many problems in which there is no actual solution or the size of the solution space is too large to accurately determine the appropriate solution. For problems like this, Orthogonality and least squares provide approximations that enable you to determine close solutions while saving on the time or the complexity of determining exact solutions.

6.1 Inner Product, Length, and Orthogonality

The inner product is also referred to as the dot product and the result of the product is a scalar value. These can be verified, but the inner product is commutative and additional properties of this operation can be found on page 333 of the text. The norm of a vector is defined as $\|\mathbf{v}\| = \sqrt{\mathbf{v} \cdot \mathbf{v}} = \sqrt{\mathbf{v}_1^2 + \mathbf{v}_2^2 + \cdots + \mathbf{v}_n^2}$ with $\|\mathbf{v}\|^2 = \mathbf{v} \cdot \mathbf{v}$. A vector whose length is one is referred to as the unit vector. By dividing a vector by its length, we are transforming a vector into a unit vector, and this process is referred to as normalizing the vector. Unit vectors can also be the negative vectors as the distance is 1. The distance between two vectors is the length of the norm of the difference between the two vectors. Two vectors are orthogonal to each other if their dot product is zero. A special case of this is with the zero vector in that every other vector is orthogonal to it. An extension of orthogonality is that two vectors are orthogonal if and only if the square of norm of their sum is equal to the sum of their norm's squared. In order to be orthogonal to a set a vector needs to be orthogonal to every vector in the set spanning the space that is under question. The orthogonal complement of the row

space of a matrix is the null space of that matrix and the orthogonal complement of the column space is the null space of the transpose of the same matrix. The orthogonal complement for a plane through the origin is a line through the origin as well. For two vectors in either 2D or 3D space, there is a nice connection between the angle between the two vectors and the inner product of the two vectors. The inner product of the two vectors is equal to the norms of those two vectors multiplied with the cosine of the angle between the two.

6.2 Orthogonal Sets

If there is a set of vectors within a region where all of the pairs of vectors within the set are orthogonal, then the set is linearly independent and the set is a basis for the region. For orthogonal sets, the scalar weights for the linear combination of the basis needed to compose the vectors in the region is given by the following formula:

$$c_j = \frac{\mathbf{y} \cdot \mathbf{u}_j}{\mathbf{u}_j \cdot \mathbf{u}_j}$$

A useful extension of this theorem is that of projection. With this, we are trying to write the vector in terms of a scalar multiple of one of the basis vectors as well as a vector that is orthogonal to the basis vector chosen. This is only achievable when the inner product of those two vectors chosen for the decomposition is zero. This leads to the previous formula being useful for determining the weight that needs to be applied to the basis vector to showcase the vector component in that direction. This is also called projecting the vector onto the basis vector. This projection is then subtracted from the vector in question to determine the other decomposition component to describe the vector. The two components then form an orthogonal set describing the vector. These theorems can be used with the basis to describe any vector in a region as the components projected onto each of the basis vectors. Vectors can form an orthonormal set if they are a orthogonal basis with each vector being of length 1. For a given $m \times n$ matrix, it has orthonormal columns if and only if the transpose of this matrix multiplied with the original matrix is the identity matrix. Orthonormal matrices that are also square have the useful property of their inverse being equal to the transpose and are referred to as orthogonal matrices. Additionally, the rows of such a matrix are also orthonormal.

6.3 Orthogonal Projections

A couple properties that stem from the findings in the previous section are that this projection is the unique vector in the space where the vector minus the projection of the vector is orthogonal to the space, and the projection is the closest vector within the space to the vector in question. These properties are the building blocks to least squares solutions to linear problems. Any vector can be written in terms of the projection for that vector to the space in question and the projection of that vector to the orthogonal space in question. It is often useful to understand how many components from the original orthogonal space are needed to span the space in question and use the rest to describe the orthogonal space. These two vectors, comprising of multiple orthogonal parts, can describe any vector in question for the given space. The number of terms in the orthogonal projection of the vector is equal to the number of terms in

the dimensionality. This makes sense as each basis vector in the vector space needed to span the space is also needed as a component to represent a given vector in this space. If a given vector is in a given vector space, the projection of this vector into this space is just the vector itself as it forms the basis. This also stems from the theorem that the projection of the vector in the subspace is the closest point to that vector for the subspace. If the basis is orthonormal, then the denominator of the projection drops out since it is 1, and this simplifies the calculation for the projection. Thus, the projection for a vector with an orthonormal basis is given by the basis multiplied with the transpose (inverse for orthonormal) of the basis and then the vector in question.

6.4 The Gram-Schmidt Process

The purpose of this section is to outline algorithms to form orthogonal or orthonormal bases for a given nonzero subspace. For a given basis of a subspace, to find the orthogonal basis of the subspace you first assign one of the vectors into the orthogonal basis, then using this vector, you project one of the original basis vectors onto this orthogonal basis vector by subtracting the parallel component out of the projection, then you use the result of that projection to project an additional basis vector into the orthogonal basis vector by subtracting out the parallel component from that projection. This process is repeated for as many basis vectors as needed to describe the subspace. An orthonormal basis is easily formed from the orthogonal basis as you just normalize each vector component by the length of the basis vector it helps to represent.

If A is an $m \times n$ matrix with linearly independent columns, then A can be factored as $A=QR$. Q corresponds to the $m \times n$ matrix whose columns form an orthonormal basis for $\text{Col } A$ and R is a $n \times n$ upper triangular invertible matrix with positive entries on the diagonal. Finding R is equivalent to multiplying the transpose of Q by A .

6.5 Least-Squares Problems

The idea behind this form of a problem is to find a solution \mathbf{x} that brings the length of the difference between the result \mathbf{b} and $A\mathbf{x}$ as low as possible or at least smaller than a given error point that is desired. The idea is to make the chosen $A\mathbf{x}$ the closest point in $\text{Col } A$ to \mathbf{b} through projections. The solution of the general least squares problem is finding the projection \mathbf{x} that satisfies the projection of \mathbf{b} into the Col space of A . These same solutions for the least squares problem correspond to the set of solutions to the normal equations which are depicted by $A^T A \mathbf{x} = A^T \mathbf{b}$. For a given solution that you are after and the transformation matrix, the least squares solution is easily found through the normal equations by either matrix algebra or row reduction. The least-squares solution is only achievable this way when the criteria next are followed:

- The matrix equation has a unique least-squares solution for each result desired.
- The columns of A are linearly independent.
- The matrix left multiplied with its transpose is invertible.

Often in these applications the point is to minimize the least-squares error. This value is the ultimate distance between the desired result and the value achieved through the transformation matrix multiplied with the found least-squares solution. When the

columns of A are orthogonal, applying the projections onto each of the components of the transformation matrix yields the least-squares solution immediately. To minimize error, a QR factorization can be used to calculate the least-squares solution by left multiplying the inverse of R by the left multiplication of the transpose of Q with the result matrix. Row reduction can instead be used with R and the left multiplication of the transpose of Q and the result to avoid inversion. If the result matrix or the projection of the result into the Col space of A are zero (orthogonal to $A\mathbf{x}$), then the solution is just solving the homogeneous equation.

6.6 Applications to Linear Models

Instead of using the x values from the raw data in order to find the least squares regression, you can take the mean-deviation form by subtracting out the mean from each of the x data points. The two columns of the design matrix will then be orthogonal and the solutions for the normal equations are simplified. Multiple regression and linear models with multiple variables can all be described with the matrix equation we have already utilized.

6.7 Inner Product Spaces

The same rules for inner products on individual vectors apply to the vector space to make up an inner product space. Functions or polynomials that are defined in the vector space are treated as vectors when they are considered as an element in a vector space. To interpret them as vectors, time steps are usually considered for the entries and make up the elements. The norm of the projection of a vector onto a subspace cannot exceed the norm of the vector itself. This aids in finding the Cauchy-Schwarz Inequality which says the absolute value of the inner product of two vectors is less than or equal to the product of the norm of those vectors. Similarly, the norm of the sum of two vectors is less than or equal to the sum of their individual norms. An definite integral of the product of two functions can be interpreted as the inner product of those functions and thus an extension of the properties found is immediate.

6.8 Applications of Inner Product Spaces

Weights can be extended to the normal equation used to determine the least-squares solution in order to vary the reliance on different data points. The weights are applied to every result matrix and transformation matrix from the previous equation. The Fourier Series can be derived from the expansion of the information learned in this chapter.