

Math 54 Review

Nicholas Lai

November 23, 2018

Singular Value Decomposition

The process of diagonalizing an $n \times n$ matrix A (fulfilling certain conditions) is a simple matter of computing

$$A = PDP^{-1}$$

Where P is a matrix with the eigenvectors of A as columns, and D is the diagonal matrix of eigenvalues with the same order as P (the eigenvalue corresponding to the eigenvector in the i th column of P must be the i th entry of D)

In general, ordinary diagonalization is not well-defined on an arbitrary $m \times n$ matrix, which motivates the Singular Value Decomposition (SVD).

Theorem:

Every $m \times n$ matrix A can be written as

$$U\Sigma V^T$$

Where U and V^T are orthogonal matrices (The columns are orthonormal to each other), and Σ is a non-negative diagonal matrix.

Computing the SVD

Finding V and Sigma

Whereas an arbitrary matrix A is not orthogonally diagonalizable in general, $A^T A$ always is, because it is symmetric. Directly from the theorem above and the properties of matrix transposes:

$$A^T A = V\Sigma^T U^T U\Sigma V^T$$

Since U is orthogonal, $U^T = U^{-1}$, so $U^T U$ is I

So we have:

$$A^T A = V\Sigma^T \Sigma V^T$$

Since Σ is diagonal, $\Sigma^T \Sigma$ is diagonal. So in fact, by orthonormally diagonalizing $A^T A$, we can obtain V^T and Σ (the entries of the diagonal matrix $\Sigma^T \Sigma$ are the squares of the values in Σ , which is non-negative)

Finding U

Whereas an arbitrary matrix A is not orthogonally diagonalizable in general, AA^T always is, because it is symmetric. Directly from the theorem above and the properties of matrix transposes:

$$AA^T = U\Sigma V^T V\Sigma^T U^T$$

Since V is orthogonal, $V^T = V^{-1}$, so $V^T V$ is I

So we have:

$$AA^T = U\Sigma\Sigma^T U^T$$

Since Σ is diagonal, $\Sigma\Sigma^T$ is diagonal. So in fact, by orthonormally diagonalizing AA^T , we can obtain U .

Notes:

1. The dimension of Σ is always $m \times n$. When putting in entries into Σ , they must be put in the i th entry corresponding to the order of your eigenvectors (similarly to normal diagonalization).
2. When you compute Σ once when obtaining V^T or U , do NOT compute it again. Use the Σ that you obtained to diagonalize the other matrix (either $A^T A$ or AA^T), as $\Sigma^T \Sigma = \Sigma \Sigma^T$. This is important not only for saving time, but also for ensuring that the ordering of vectors in V^T and U agree. The entries of $\Sigma^T \Sigma$ and $\Sigma \Sigma^T$ are the eigenvalues of $A^T A$ and AA^T .
3. The dimension of U is $m \times m$.
4. The dimension of V^T is $n \times n$.
5. You MUST orthonormally diagonalize $A^T A$ and AA^T in order to calculate the SVD, or else the above procedure is NOT valid. Remember that orthonormal diagonalization is the same as diagonalization, except that you apply Gram-Schmit to each eigenspace basis to obtain an orthonormal basis of eigenvectors.
6. Eigenvectors in different eigenspaces are naturally orthogonal (If eigenvectors correspond to different eigenvalues, they are orthogonal). You only need to apply Gram-Schmit to vectors of a given eigenvalue if there is more than one for that eigenvalue, and you only need to perform Gram-Schmit on those vectors, not all the eigenvectors. If not, all you need to do is normalize the vector. This is an important time-saving trick because Gram-Schmit increases in complexity the more vectors you are trying to orthonormalize, so keep those groups as small as possible!

An Example for Note 6

http://www.math.jhu.edu/~hhaosu/Teaching/0203SLA/Notes-Ch8_1.pdf

Notice how it suffices to get orthonormal bases for each individual eigenspace, as the bases will naturally be orthogonal to each other.

Orthogonal Projection

Given a vector \vec{v} in the vector space V , and a subspace S of V , \vec{v} can be written as the sum of two unique vectors \vec{a} and \vec{b} , where \vec{a} is in the subspace S and \vec{b} is in the orthogonal complement of S .

We denote \vec{a} as the projection of \vec{v} on S .

Calculating Projections

The projection of a vector \vec{v} onto another vector \vec{s} :

$$\text{proj}_{[\vec{s}]}(\vec{v}) = \frac{\vec{v} \cdot \vec{s}}{\vec{s} \cdot \vec{s}} \vec{s}$$

Where \cdot is the dot product of vectors. The orthogonal component of \vec{v} to \vec{s} can be obtained by subtracting the projection from the original vector \vec{v}

We can calculate the projection of a vector \vec{v} onto an arbitrary subspace S with a basis matrix denoted A :

$$\text{proj}_{[S]}(\vec{v}) = A(A^T A)^{-1} A^T \vec{v}$$

Basis

A basis of a vector space is a linearly independent set of vectors that spans the entire vector space.

Important Properties of a Basis

1. A Basis is not unique, in general.
2. A square matrix whose column vectors form a basis is invertible, and has all the properties that the Invertible Matrix Theorem entails.
3. The matrix transformation defined by A square matrix whose column vectors form a basis is onto and one-to-one (surjective and injective).
4. A basis of \mathbb{R}^n has n elements
5. All elements in the vector space spanned by a basis have a unique representation as a linear combination of elements in the basis.

Checking if a set of vectors is a basis

The simplest way to do this is to see if the matrix representation of the basis is full rank and nullity. This can be checked by row-reducing the matrix.

More formally, you can also use the definition of linear independence (show that the only solution to the homogeneous equation is the vector of all zeroes) and show that all vectors in the vector space can be written as a linear combination of basis elements. This is usually not required until Math 110.

Transformations

A transformation is an operation that takes elements from one vector space and outputs elements of another vector space, not necessarily the same.

A transformation f is linear if it satisfies:

$$f(c\vec{u} + \vec{v}) = cf(\vec{u}) + f(\vec{v})$$

Theorem:

Matrix transformations are linear transformations.

So, this means that if you can represent a function with a matrix, you have proven that the function is linear, and can use linearity properties.

Complex Eigenvalues

Theorem:

For real-valued matrices, if a matrix has a complex eigenvalue, its complex conjugate is also an eigenvalue.

An example for solving for complex eigenspaces: https://www.youtube.com/watch?v=d9neNf8_b0I

Symmetric Matrices

A Matrix A is symmetric if $A = A^T$. From this fact directly follows that $A^{-1}A^T = I$

Theorem:

If a real-valued matrix is symmetric, it has only real eigenvalues.

Theorem:

Symmetric matrices are orthogonally diagonalizable

This theorem underpins the SVD procedure that we performed above. $A^T A$ and AA^T are symmetric for any matrix, and therefore can always be orthogonally diagonalized.