Diagonalization & Singular-Value Decomposition

Zvi Rosen
Department of Mathematics

October 28, 2016

Idea

Use eigenvalues and eigenvectors to factor a matrix into a more useful product of matrices. (Diagonalization)

Generalize this concept to non-diagonalizable matrices.

Diagonalization

We know how to compute eigenvalues and eigenvectors, whether using MATLAB or by the characteristic polynomial. Fix a (diagonalizable) matrix A. Let D be a diagonal matrix whose diagonal entries are the eigenvalues $\lambda_1, \ldots, \lambda_n$ of A, and P be the matrix whose columns are the corresponding eigenvectors $x_1, \ldots, x_n$, which form a basis of the vector space. Then:

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

Why? Consider how this decomposition acts on the eigenvectors.

Non-diagonalizable matrices

Not every matrix is diagonalizable.
Consider the matrix

$$\begin{matrix} 1 & 1 \\ 0 & 1 \end{matrix}$$

This does not have a set of eigenvectors that span the space.
How can we tell?

Application: Matrix Powers

In many applications, we want to compute A^M without doing $M - 1$ matrix multiplications. If we have a diagonalization $A = PDP^{-1}$, then

$$A^M = (PDP^{-1})(PDP^{-1}) \cdots (PDP^{-1})$$

$$\implies A^M = PD(P^{-1}P)D(P^{-1} \cdots P)DP^{-1}) = PD^M P^{-1}.$$

Power of a diagonal matrix D^M is the diagonal matrix with entries raised to the power.

Singular Value Decomposition

Generalizes the notion of diagonalization for arbitrary $m \times n$ matrices.

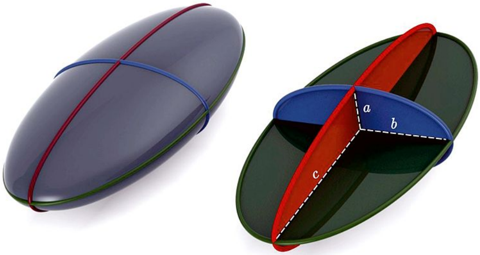
$A = U\Sigma V^*$ is a decomposition with U, V orthogonal matrices – corresponding to rotations – and Σ is a diagonal matrix ordered from largest to smallest entries – corresponding to stretching. The entries of Σ are called singular values of the matrix. They are eigenvalues of $MM*$ and $M * M$.

SVD: Introduction

Strongly recommended piece on SVD:
http://www.ams.org/samplings/feature-column/fcarc-svd.

Low rank approximation

If we throw away smaller singular values of the matrix, the matrix keeps doing its large-scaling operations, but instead of shrinking some vectors, it just collapses them to zero.



Imagine that instead of returning a vector in the larger ellipsoid at left, it returns a vector in the oval. (Image from Wikipedia)

Applications: Image Processing

An image with an $m \times n$ grid of grayscale pixels can be written as an $m \times n$ matrix with entries between 0 and 255.

Applications: Image Processing

For instance, consider this image:

This has 700×700 pixels, almost all black or white, with some grayscale.