CS205 – Class 5

***Covered in class:** 1, 3, 4, 5.*
***Reading:** Heath Chapter 4.*

Eigenvalues / Eigenvectors Continued

1. The *Power Method* allows one to compute the largest eigenvalue and eigenvector. Starting from a nonzero vector x_0, iterate with $x_{k+1} = Ax_k$.
    a. To see why this works, assume that x_0 is a linear combination of eigenvectors $x_0 = \sum_i \alpha_i u_i$ where the u_i are the eigenvectors of A. Then $x_k = Ax_{k-1} = A^2 x_{k-2} = \cdots = A^k x_0$ and so $x_k = A^k x_0 = A^k \sum_i \alpha_i u_i = \sum_i \alpha_i A^k u_i = \sum_i \alpha_i \lambda_i^k u_i$. Now assuming that the largest eigenvalue is λ_1, we can write $x_k = \alpha_1 \lambda_1^k u_1 + \sum_{i=2} \alpha_i \lambda_i^k u_i = \lambda_1^k \left(\alpha_1 u_1 + \sum_{i=2} \alpha_i \left(\lambda_i / \lambda_1 \right)^k u_i \right)$ and note that the second term vanishes as $k \to \infty$ since $|\lambda_i / \lambda_1| < 1$. Thus as $k \to \infty$, $x_k \to \lambda_1^k \alpha_1 u_1$. Moreover $\left(x_k \right)_j / \left(x_{k-1} \right)_j \to \lambda_1$ for any component j of x.
    b. If the starting vector $x_0 = \sum_i \alpha_i u_i$ happens to have $\alpha_i = 0$ for the largest eigenvalue, the method fails.
    c. For a real matrix and real x_0, one can never get complex numbers.
    d. The largest eigenvalue may be repeated, in which case the final vector may be a linear combination of the true eigenvectors.
    e. After every iteration, x_k can be renormalized to stop x_k from growing too large.
    f. Shifts can be used to accelerate convergence.
    g. Inverse iteration can be used to find the smallest eigenvalue. This relies on the fact that the eigenvalues of A^{-1} are the reciprocals of those of A. Thus, the largest eigenvalue of A^{-1} is the smallest eigenvalue of A.
    h. *Deflation* is a method to remove an eigenvalue from a matrix A once it has been computed. Then the resulting matrix can be analyzed to compute the next largest eigenvalue, etc.
2. If $Ax = \lambda x$, then one can form the Rayleigh Quotient $\lambda = \frac{x^T Ax}{x^T x}$. This is used in a variety of methods for computing eigenvalues.

Singular Value Decomposition (SVD)

3. The Singular Value Decomposition is an eigenvalue-like decomposition for rectangular $m \times n$ matrices. It has the form $A = U \sum V^T$ where U is an $m \times m$ orthogonal matrix, *V* is an $n \times n$ orthogonal matrix, and $\sum$ is an $m \times n$ diagonal matrix with positive diagonal entries that are called the *singular values* of A. The columns of *U* and *V* are the *singular vectors*.
    a. Introduced and rediscovered many times: Beltrami in 1873, Jordan in 1875, Sylvester in 1889, Autonne in 1913, Eckart and Young in 1936.
    b. Pearson introduced principle component analysis (PCA) in 1901. It uses SVD.
    c. Numerical work by Chan, Businger, Golub, Kahan, etc.

4. The singular value decomposition of $A = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \\ 10 & 11 & 12 \end{bmatrix}$ is given by

$$\begin{bmatrix} .141 & .825 & -.420 & -.351 \\ .344 & .426 & .298 & .782 \\ .547 & .028 & .664 & -.509 \\ .750 & -.371 & -.542 & .079 \end{bmatrix} \begin{bmatrix} 25.5 & 0 & 0 \\ 0 & 1.29 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} .504 & .574 & .644 \\ -.761 & -.057 & .646 \\ .408 & -.816 & .408 \end{bmatrix}.$$

   a. The singular values are 25.5, 1.29, and 0. The singular value of 0 indicates that the matrix is rank deficient. However, even a “small” singular value could indicate a “zero” and a rank deficient matrix.

5. The singular values of A are the non-negative square roots of the eigenvalues of the symmetric positive semi-definite $A^T A$ (and also AA^T), and the columns of U and V are the orthonormal eigenvectors of AA^T and $A^T A$ respectively. (Note the strong connection to the normal equations and least squares problems).
6. The condition number of a matrix A with respect to the Euclidean norm is $\sigma_{\max} / \sigma_{\min}$.
   a. For a square matrix, the condition number measures the closeness to singularity. For a rectangular matrix, the condition number measures the closeness to rank deficiency.
7. The rank of a matrix is equal to the number of nonzero singular values that it has. However, if values are ”close” to “zero” then the condition number $\sigma_{\max} / \sigma_{\min}$ can be very high essentially making these numbers “zero” as far as rank is concerned.
8. The columns of V corresponding to “zero” singular values form an orthonormal basis for the null space of A.
   a. The remaining columns of V form an orthonormal basis for the space perpendicular to the null space of A.
9. The columns of U corresponding to the “nonzero” singular values form an orthonormal basis for the range of A.
   a. The remaining columns of U form an orthonormal basis for the space perpendicular to the range of A.
10. The columns of V corresponding to zero columns of Σ and the columns of U corresponding to zero rows of Σ along with those zero columns and rows can then be omitted without changing their product.
    a. Applying this to the SVD of A from part 4 gives us the new reduced SVD,

$$\begin{bmatrix} .141 & .825 \\ .344 & .426 \\ .547 & .028 \\ .750 & -.371 \end{bmatrix} \begin{bmatrix} 25.5 & 0 \\ 0 & 1.29 \end{bmatrix} \begin{bmatrix} .504 & .574 & .644 \\ -.761 & -.057 & .646 \end{bmatrix}$$