

## CS205 – Class 5

*Covered in class: 1, 3, 4, 5.*

*Reading: Heath Chapter 4.*

### **Eigenvalues / Eigenvectors Continued**

1. Idea is to compute eigenvalues using matrix form because it is in general easier than solving a  $n$ -degree polynomial.
  - a. Must be careful though as we need to preserve least squares solution. i.e. Gaussian-Elimination will not work.
  - b. Use a similarity transform that produces a new matrix with “same” eigenvalues/eigenvectors.
2. Formally, a matrix  $A$  is said to be **similar** to a matrix  $B$ , if  $B = T^{-1}AT$  for a nonsingular matrix  $T$ .
  - a. If  $A$  and  $B$  are similar, then they have the same eigenvalues.  $By = \lambda y$  or  $T^{-1}ATy = \lambda y$  or  $A(Ty) = \lambda(Ty)$ . Note that the eigenvectors of  $A$  are  $Ty$  where  $y$  are the eigenvectors of  $B$  (this formalizes the notion of sameness from 3(b)).
  - b. If the matrix  $A$  has distinct eigenvalues (no repeated eigenvalues), then similarity transforms can be used to put it into diagonal form where the eigenvalues can be read from the diagonal and the eigenvectors are the columns of the identity matrix. Then the eigenvectors of  $A$  are  $T$  times the columns of the identity matrix, i.e. the columns of  $T$ .
  - c. If  $A$  is real and symmetric, an orthogonal  $T$  can be used to put  $A$  into diagonal form. Moreover, the eigenvalues are real valued.
  - d. If  $A$  is complex and Hermitian, a unitary  $T$  can be used to put  $A$  into diagonal form. Moreover, the eigenvalues are real valued.
  - e. If  $A$  is normal, a unitary  $T$  can be used to put  $A$  into diagonal form.
  - f. Any matrix can be put into upper triangular, Shur form, with a unitary  $T$ . Then the eigenvalues can be read off the diagonal of the matrix.
  - g. Any matrix can be put into *Jordan form* where the eigenvalues are on the diagonal, and off diagonal elements only occur on the band above the diagonal and only for defective

eigenvalues. For example  $\begin{bmatrix} 2 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 3 & 1 \\ 0 & 0 & 0 & 3 \end{bmatrix}$  is in Jordan form where both 2 and 3 are

repeated eigenvalues, but only 3 is defective. The number of Jordan blocks = number of eigenvalues.

3. The condition number for an eigenvalue problem is defined by  $1/|y^H x|$  where  $x$  and  $y$  are the normalized right and left eigenvectors.
- If  $x$  and  $y$  are real valued, then  $y^H x = y^T x$ . Note that  $y^T x = |y||x|\cos\theta = \cos\theta$  where  $\theta$  is the angle between the eigenvectors.
  - For symmetric and Hermitian matrices, the left and right eigenvectors are the same so the condition number is 1.
  - Eigenvalues are well conditioned for normal matrices.
  - Multiple or “close” eigenvalues can be poorly conditioned, especially if they are defective or “close” to being defective.
  - Scaling by a diagonal similarity transform – called *balancing* – can improve the condition number of an eigenvalue problem.
4. Numerical methods for finding eigenvalues and eigenvectors.
- Would like a technique that solves them one at a time starting with the largest in magnitude. Similarity transform is less desirable because it gives us all of them even we need only a few.
  - Using the characteristic polynomial is a bad idea since the coefficients are ill-conditioned. Moreover, one should set up a matrix and compute eigenvalues in order to find the roots of a polynomial equation.
  - QR iteration. Initially set  $A_0 = A$  and then iterate  $A_1, A_2, \dots$ . For each  $k$ , compute the QR factorization  $A_k = Q_k R_k$ , and then define  $A_{k+1} = R_k Q_k = Q_k^H Q_k R_k Q_k = Q_k^H A_k Q_k$ .
    - If the eigenvalues are all distinct, then the  $A_k$  converge to a triangular matrix. Moreover, if  $A$  is symmetric, the  $A_k$  converge to a diagonal matrix.
    - This convergence can be accelerated by using *shifts* of the form  $A_k - \sigma_k I = Q_k R_k$  and  $A_{k+1} = R_k Q_k + \sigma_k I$  where  $\sigma_k$  is a rough approximation to an eigenvalue. Initially one can use, for example, the number in the lower right hand corner of the matrix.
  - The *Power Method* allows one to compute the largest eigenvalue and eigenvector. Starting from a nonzero vector  $x_0$ , iterate with  $x_{k+1} = Ax_k$ .
    - 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  $x_k = Ax_{k-1} = A^2 x_{k-2} = \dots = A^k x_0$  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  $\lambda_1$ , we can write  $x_k = \alpha_1 \lambda_1^k u_1 + \sum_{i=2} \alpha_i \lambda_i^k u_i = \lambda_1^k \left( \alpha_1 u_1 + \sum_{i=2} \alpha_i (\lambda_i / \lambda_1)^k u_i \right)$  and note that the second term vanishes as  $k \rightarrow \infty$  since  $|\lambda_i / \lambda_1| < 1$ . Thus as  $k \rightarrow \infty$ ,  $x_k \rightarrow \lambda_1^k \alpha_1 u_1$ . Moreover  $(x_k)_j / (x_{k-1})_j \rightarrow \lambda_1$  for any component  $j$  of  $x$ .

- ii. 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.
  - iii. For a real matrix and real  $x_0$ , one can never get complex numbers.
  - iv. The largest eigenvalue may be repeated, in which case the final vector may be a linear combination of the true eigenvectors.
  - v. After every iteration,  $x_k$  can be renormalized to stop  $x_k$  from growing too large.
  - vi. Shifts can be used to accelerate convergence.
  - vii. 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  $A$ .
  - viii. *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.
- e. 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.