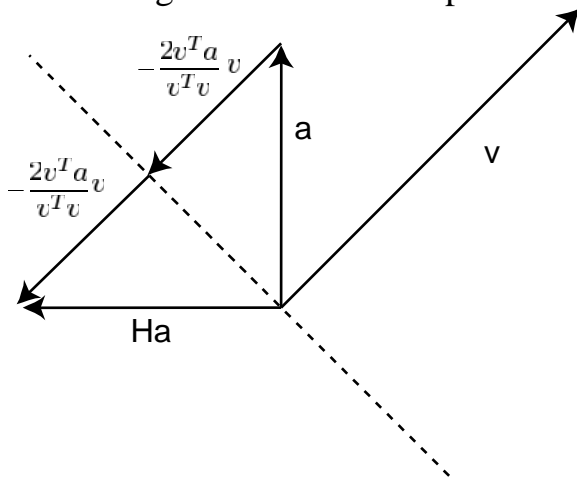


CS205 - Class 4

QR Continued (Readings Heath pp120-137)

1. Transition to Householder. So far, we've seen two different ways of finding a least squares solution, namely the normal equations and the (modified) Gram-Schmidt method. Using the normal equations squares the condition number of the A matrix, which could be bad to begin with. The Gram-Schmidt method does not suffer from the same numerical problems, but can still be unstable. If the columns of A are "nearly linearly-dependent," then this method can suffer from numerical instabilities. The QR decomposition is exactly what the Gram-Schmidt procedure does couched in the language of matrix factorization. There is an algorithm to performs the QR factorization that does not suffer from the numerical instabilities mentioned above and this is the one that is quite often used by practitioners.
2. A **Householder transform** is defined by $H = I - \frac{2}{v^T v} v v^T$ for some vector $v \neq 0$. Note that $H = H^T = H^{-1}$, and thus H is orthogonal.
 - a. For a vector a , we define H_k using $v_k = \hat{a} + S(a_k) \|\hat{a}\|_2 e_k$ where $\hat{a} = (0, \dots, 0, a_k, \dots, a_m)^T$ and $S(a_k) = \pm 1$ is the sign function. Then $H_k a$ zeroes out the entries of a below a_k .
 - b. Let $a = \begin{bmatrix} 2 \\ 1 \\ 2 \end{bmatrix}$. Then $v_1 = \hat{a} + S(a_1) \|\hat{a}\|_2 e_1 = \begin{bmatrix} 2 \\ 1 \\ 2 \end{bmatrix} + S(2) \sqrt{9} \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 5 \\ 1 \\ 2 \end{bmatrix}$.
 - c. Note that $H a = a - \frac{2 v^T a}{v^T v} v$ so we never need to form H explicitly, but instead only need to find the vector v .
 - d. $H_1 \begin{bmatrix} 2 \\ 1 \\ 2 \end{bmatrix} = \begin{bmatrix} 2 \\ 1 \\ 2 \end{bmatrix} - \frac{2(5,1,2)(2,1,2)^T}{(5,1,2)(5,1,2)^T} \begin{bmatrix} 5 \\ 1 \\ 2 \end{bmatrix} = \begin{bmatrix} 2 \\ 1 \\ 2 \end{bmatrix} - \frac{2 \times 15}{30} \begin{bmatrix} 5 \\ 1 \\ 2 \end{bmatrix} = \begin{bmatrix} -3 \\ 0 \\ 0 \end{bmatrix}$.
 - e. In order to use the householder transform to compute the QR factorization of a matrix A, we define H_k using $v_k = \hat{a}^{(k)} + S(a_{kk}) \|\hat{a}^{(k)}\|_2 e_k$ where $\hat{a}^{(k)} = (0, \dots, 0, a_{kk}, \dots, a_{mk})^T$. When applying H_k to the matrix A, $H_k A$ is obtained by applying H_k to every column of A.
 - f. Overall, we obtain $H_n \dots H_1 A = \begin{bmatrix} R \\ 0 \end{bmatrix}$. Then applying $Q = H_1^T \dots H_n^T$ (all the H's are orthogonal and so is their product) to both sides of the equation results in $A = Q \begin{bmatrix} R \\ 0 \end{bmatrix}$.

- g. Thus, to solve $Ax=b$, we write $Q \begin{bmatrix} R \\ 0 \end{bmatrix} x = b$, and $\begin{bmatrix} R \\ 0 \end{bmatrix} x = Q^T b = H_n \cdots H_1 b = \hat{b}$, and then solve $\begin{bmatrix} R \\ 0 \end{bmatrix} x = \hat{b}$ with an upper triangular solve. That is, starting with $Ax=b$, we simply apply $Q^T = H_n \cdots H_1$ to each side, one H_k at a time to obtain $\begin{bmatrix} R \\ 0 \end{bmatrix} x = \hat{b}$, and then solve this upper triangular problem.
- h. Again, thinking geometrically, we can see that the Householder transformations are actually reflections about the hyperplane orthogonal to v . This can be seen by H subtracting off twice the component of x in the direction of v .



- i. We motivated this entire QR discussion by solving an overdetermined system of linear equations. Often there will be no solution. There are too many constraints, but we can still strive for the best solution, the least-squares solution. We have shown that factorization preserves the least squares solution while Gaussian elimination fails to do so. The intuition behind this fact is that the matrix Q shuffles the equations in such a way that the least squares solution exactly solves the first n equations, while leaving the minimum amount of slack in the remaining $m-n$ equations. Gaussian elimination fails to shuffle the equations in a way that keeps the norm of the residual at a minimum.

Eigenvalues and Eigenvectors (Readings Heath pp157-160)

3. For an $n \times n$ matrix A , $Ax = \lambda x$ is the standard eigenvalue problem where λ is an eigenvalue and x is a left eigenvector.
- The right eigenvectors y are defined by $y^T A = \lambda y^T$. If y is a right eigenvector of A , then it is a left eigenvector of A^T , since $A^T y = \lambda y$.
 - Usually we refer to “left” eigenvectors simply as eigenvectors while still referring to “right” eigenvectors as right eigenvectors.

- c. The set of all eigenvalues of A is called the *spectrum* of A , and the *spectral radius* of A is the magnitude of its largest magnitude eigenvalue. $\rho(A) = \max_i |\lambda_i|$.
- d. An eigenvector is a direction along which the action of a matrix is rather simple. The matrix merely expands or contracts vectors in that direction (according to the eigenvalue) leaving the direction unchanged.
- e. For diagonal matrices, the eigenvalues are on the diagonal, and the eigenvectors are the columns of the identity matrix. For example, $\begin{bmatrix} 2 & 0 \\ 0 & 3 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} = 2 \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ and $\begin{bmatrix} 2 & 0 \\ 0 & 3 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \end{bmatrix} = 3 \begin{bmatrix} 0 \\ 1 \end{bmatrix}$.
- f. For upper triangular and lower triangular matrices, the eigenvalues appear on the diagonal. For example, $\begin{bmatrix} 2 & 1 \\ 0 & 3 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} = 2 \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ and $\begin{bmatrix} 2 & 1 \\ 0 & 3 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \end{bmatrix} = 3 \begin{bmatrix} 1 \\ 1 \end{bmatrix}$.

4. Recall an eigenvalue is a scalar that satisfies $Ax = \lambda x$

- a. Upper triangular matrices have eigenvalues on diagonal
- b. Symmetric matrices have real eigenvalues
- c. A symmetric matrix is guaranteed to have real eigenvalues, i.e. $\lambda \in \mathbb{R}$, while nonsymmetric matrices can have complex eigenvalues. For example $\begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \begin{bmatrix} 1 \\ i \end{bmatrix} = i \begin{bmatrix} 1 \\ i \end{bmatrix}$.

It is easy to see why symmetric (or Hermitian, in general) matrices have real eigenvalues. Let $Ax = \lambda x$, then $\lambda \langle x, x \rangle = \lambda x^H x = x^H \lambda x = x^H Ax = x^H A^H x = \bar{\lambda} x^H x = \bar{\lambda} \langle x, x \rangle$

- d. Eigenvectors can be arbitrarily scaled by a constant, for example $\begin{bmatrix} 2 & 1 \\ 0 & 3 \end{bmatrix} \begin{bmatrix} s \\ s \end{bmatrix} = 3 \begin{bmatrix} s \\ s \end{bmatrix}$ for any s . Thus, we usually require that eigenvectors be normalized (usually in the L_2 norm). Thus, $\begin{bmatrix} 2 & 1 \\ 0 & 3 \end{bmatrix} \begin{bmatrix} \sqrt{2}/2 \\ \sqrt{2}/2 \end{bmatrix} = 3 \begin{bmatrix} \sqrt{2}/2 \\ \sqrt{2}/2 \end{bmatrix}$ gives the eigenvector in standard form.
- e. $Ax = \lambda x$ can be written equivalently as $(A - \lambda I)x = 0$ and there is a nontrivial solution x to this problem when $A - \lambda I$ is NOT invertible, or singular, or $\det(A - \lambda I) = 0$.
 - i. Note that $\det(A - \lambda I)$ is an n -th degree polynomial and we refer to it as the *characteristic polynomial* of A . The roots of the *characteristic equation* $\det(A - \lambda I) = 0$ are the eigenvalues of A .
 - ii. $\det\left(\begin{bmatrix} 2 & 1 \\ 0 & 3 \end{bmatrix} - \lambda \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}\right) = \det\left(\begin{bmatrix} 2-\lambda & 1 \\ 0 & 3-\lambda \end{bmatrix}\right) = (2-\lambda)(3-\lambda) = 0$. Thus the matrix $\begin{bmatrix} 2 & 1 \\ 0 & 3 \end{bmatrix}$ has $\lambda = 2, 3$ as eigenvalues.
 - iii. An $n \times n$ matrix A has an n -th degree characteristic polynomial with n roots, and thus n eigenvalues. However, there may be multiple roots or complex roots.
 - 1. In the repeated eigenvalue case, if there are fewer linearly independent eigenvectors than repeated roots, the eigenvalue and the matrix is said to be *defective*.

2. $\begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} = 2 \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ and $\begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \end{bmatrix} = 2 \begin{bmatrix} 0 \\ 1 \end{bmatrix}$ so the multiple eigenvalue of 2 has two linearly independent eigenvectors.
3. $\begin{bmatrix} 2 & 1 \\ 0 & 2 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} = 2 \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ is the only eigenvector relationship for the multiple eigenvalue 2 here, and thus this matrix is defective.

5. Special matrices

- Symmetric* matrices have $A^T = A$.
- Hermitian* matrices have $A^H = A$ where the “H” superscript indicates the complex conjugate of the transpose. Thus, for matrices with real values only, $A^H = A^T$, i.e. “H” corresponds to simple transposition
- Orthogonal* matrices have $A^T A = A A^T = I$, i.e. the columns are orthonormal.
- Unitary* matrices have $A^H A = A A^H = I$.
- Normal* matrices have $A^H A = A A^H$
- Positive Definite* matrixes have $x^T A x > 0$ for all non-zero x .