HW6

1. Consider the Conjugate Gradient algorithm in class. The search directions are obtained as follows

$$p^{(0)} = r^{(0)} = b - Ax^{(0)}.$$

For later directions, one construct

$$p^{(k+1)} = r^{(k+1)} + s_k p^{(k)}$$

One chooses s_k such that $p^{(k+1)}$ is A-orthogonal to $p^{(k)}$. Fill in the details (and repeat) in the proof in lecture notes to

- Show that p^i's are A-orthogonal search directions.
- Show that the formula for s can be simplified to the following.

$$s_k = \frac{|r^{(k+1)}|^2}{|r^{(k)}|^2}.$$

Find the lecture notes for the proof.

2. Find the singular value decomposition of

$$A = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 1 & 1 \\ 0 & 1 & 0 \end{pmatrix}$$

3. Prove the following using singular value decomposition

- Show that the 2-norm of A is the largest singular value and the Frobenius norm is the square root of the sum of sqaures of all singular values.

Consider the SVD:

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

where U is an $m \times m$ orthogonal matrix, V is an $n \times n$ orthogonal matrix, and $\Sigma_{ij} = \sigma_i \delta_{ij}$ with $\sigma_1 \geq \sigma_2 \geq \cdots \sigma_{\min(m,n)} \geq 0$.

Let $D = \Sigma^T \Sigma = (\sigma_i^2 \delta_{ij})_{n\times n}$ and it is thus an $n \times n$ square matrix. We have assumed $\sigma_i = 0$ for $i \geq \min(m, n)$. The 2-norm is given by

$$\|A\|_2 = \sqrt{\rho(A^T A)} = \sqrt{\rho(V\Sigma^T U^T U\Sigma V^T)} = \sqrt{\rho(VDV^T)} = \sqrt{\rho(D)} = \sigma_1.$$

Above, we have used the fact $\lambda(VDV^T) = \lambda(D)$.
For the Frobenius norm, we have

$$\|A\|_F = \sqrt{\sum_{i,j} a_{ij}^2} = \sqrt{\mathrm{tr}(A^T A)} = \sqrt{\mathrm{tr}(VDV^T)}$$
$$= \sqrt{\mathrm{tr}(DV^T V)} = \sqrt{\mathrm{tr}(D)} = \sqrt{\sum_i \sigma_i^2}.$$

- Show that AA^T and $A^T A$ have the same 2 norm.
  Note that AA^T is not the transpose of $A^T A$!
  Using the same SVD as above, we have

  $$AA^T = U\Sigma\Sigma^T U^T, \quad A^T A = VDV^T.$$

  Since AA^T is real symmetric, the 2-norms is the largest eigenvalue, or the spectral radius, given by $\rho(\Sigma\Sigma^T) = \sigma_1^2$. Similarly, $\|A^T A\|_2 = \rho(D) = \sigma_1^2$.

4. Write a code for power method and apply it to the following matrix to find the leading eigen-pair:

$$\begin{pmatrix} -4 & 14 & 0 \\ -5 & 13 & 0 \\ -1 & 0 & 2 \end{pmatrix}$$

5. Suppose matrix A is a real matrix with eigenvalues $|\lambda_1| > |\lambda_2| \geq \cdots \geq |\lambda_n|$. Show that the power method still works even if the matrix A does not have n eigenvectors (you may consider the Jordan canonical form).

   According to power method,

   $$x^{(k)} = \frac{1}{\mu_k} A^k x_0,$$

   where μ_k is some scalar to make the $\|\cdot\|_\infty$ norm to be 1. Clearly, the point is to check how $A^k x_0$ behaves. We want to say that the component corresponding ot λ_1 dominates.

   We use the Jordan canonical form:

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

According to the assumption, we know

$$J = \begin{pmatrix} \lambda_1 & & & \\ & J_{\mu_2} & & \\ & & \cdots & \\ & & & J_{\mu_m} \end{pmatrix}$$

where J_{μ_k} are some Jordan blocks. Hence, the first column of P is the eigenvector corresponding to λ_1. Though there are no enough eigenvectors, we use the columns of P, $\{v_1, \cdots, v_n\}$, as the basis (the generalization of eigenvectors):

$$A^k x_0 = PJ^kP^{-1}(\sum_{i=1}^n \alpha_i v_i)$$

Since $I = P^{-1}P = P^{-1}[v_1, \cdots, v_n]$, we find $P^{-1}v_i = e_i$ (the vector with ith component being 1 and others being zero.) Hence,

$$A^k x_0 = \sum_{i=1}^n \alpha_i (PJ^k)_i = \alpha_1 \lambda_1^k v_1 + \sum_{i=2}^n \alpha_i w_{i,k}$$

where B_i means the ith column of B, and $w_{i,k} = (PJ^k)_i = P(J^k)_i$. The ith column of J^k corresponds to one Jordan block: $J^k_{\mu_\ell}$ for some ℓ. Direct computation shows

$$J^k_{\mu_\ell} = (\mu_\ell I + F)^k = \sum_{j=0}^{\min(n,k)} C^j_k \mu_\ell^{k-j} F^j.$$

This shows that

$$|w_{ik}| \leq C \sum_{j=0}^n |k|^j |\mu_\ell|^{k-j} \ll |\lambda_1|^k.$$

Note that μ_ℓ is λ_i. In other words,

$$\frac{|A^k x_0 - \alpha_1 \lambda_1^k v_1|}{|\lambda_1|^k} \to 0.$$

The eigenvector then converges. Hence, the power method works.

6. This is about Householder transform

(a) The Householder transform

$$P = I - 2w^T w$$

is the identity matrix perturbed a rank 1 matrix. Show that $n-1$ eigenvalues are unchanged while one eigenvalue is changed to be -1. (Hint: use the geometric meaning of the transform. Try to identify the eigenvectors using the geometric meaning directly)

For identity matrix, all eigenvalues are 1. With the rank 1 perturbation $-2ww^T$, one eigenvalue is changed. To figure out how this is changed, we check the geometric meaning.

According to the geometric meaning, w is an eigenvector with egienvalue -1, while all the nonzero vectors perpendicular to w are eigenvectors with eigenvalue 1. Since the space perpendicular to w has dimension $n-1$, we thus can find $n-1$ independent such eigenvectors. Hence, all the n eigenvectors are found.

(b) Let H_1 be a Householder transform in $\mathbb{R}^{n-1}$. Show that transform

$$Q_1 = \begin{pmatrix} 1 & 0 \\ 0 & H_1 \end{pmatrix}$$

is also a Householder transform.

Let $H_1 = I_{n-1} - 2w_1 w_1^T$ for some $w_1 \in \mathbb{R}^{n-1}$ with $|w_1| = 1$. According to the formula, the matrix Q_1 is doing the following: the first component of x is unchanged, while for the other $n-1$ components, it is doing the reflection by H_1. Geometrically, it should be a refection in $\mathbb{R}^n$ as well. This motivates us to consider

$$w = \begin{pmatrix} 0 \\ w_1 \end{pmatrix}.$$

Hence, we guess:

$$Q_1 = I_n - 2ww^T.$$

In fact, one can check this easily:

$$I_n - 2ww^T = \begin{pmatrix} 1 & 0 \\ 0 & I_{n-1} \end{pmatrix} - 2\begin{pmatrix} 0 & 0 \\ 0 & w_1 w_1^T \end{pmatrix} = Q_1.$$

Lastly, it is easy to see $|w| = |w_1| = 1$. The claim is then proved.

7. Give a 2×2 matrix such that the diagonal elements of A_k in the QR algorithm do not converge to the eigenvalues.

   You can construct a matrix with complex eigenvalues. Other examples are also possible, for example, many of you constructed $A = [0, 1; 1, 0]$

8. Apply QR algorithm (you may code) to find all eigenvalues of the following matrix (with accuracy 10^{-4})

$$\begin{pmatrix} 3 & 1 & 0 \\ 1 & 4 & 2 \\ 0 & 2 & 3 \end{pmatrix}$$