Homework Assignment 1 – Math 118, Winter 2021

Dhruv Chakraborty

January 26, 2021

Problem 1

Suppose that $A \in \mathbb{R}^{m\times n}$, $\mathbf{v} \in \mathbb{R}^n$ and $B \in \mathbb{R}^{n\times r}$. Determine the number of floating point operations (*ie* additions, multiplications and divisions) required to compute $A\mathbf{v}$ and AB.

Solution: To compute $A\mathbf{v}$, we calculate the dot product $\mathbf{a}_i^\top \mathbf{b}$ m times, where $\mathbf{a}_i$ is the i-th row of A. Computing each dot product requires n multiplications and $n-1$ additions, so computing $A\mathbf{v}$ takes $2mn - m$ floating point operations.

When computing AB, we repeatedly calculate $A\mathbf{b}_i$ to find the i-th column of AB where $\mathbf{b}_i$ is the i-th column of B. Since B has r columns, we do this r times, which means calculating AB requires $2mnr - mr$ floating point operations.

Problem 2

Let $\mathbf{v} = [1, 0, 5, -3]^\top$ and $\mathbf{w} = [-2, 4, 5, 1]^\top$.

1. Compute $\|\mathbf{v}\|_1$, $\|\mathbf{w}\|_3$ and $\|\mathbf{v}\|_\infty$.
2. Verify that $\langle \mathbf{v}, \mathbf{w} \rangle \leq \|\mathbf{v}\|_2 \|\mathbf{w}\|_2$ by computing both sides.

Solution:

1. - $\|\mathbf{v}\|_1 = \sum_{i=1}^4 |v_i| = 9$
   - $\|\mathbf{w}\|_3 = \left(\sum_{i=1}^4 |w_i|^3 \right)^{1/3} = 198^{1/3} \approx 5.828$
   - $\|\mathbf{v}\|_\infty = \max_i |v_i| = 5$
2. $\langle \mathbf{v}, \mathbf{w} \rangle = \mathbf{v}^\top \mathbf{w} = 20 \leq \sqrt{1610} = \sqrt{35} \cdot \sqrt{46} = \left(\sum_{i=1}^4 v_i^2 \right)^{1/2} \left(\sum_{i=1}^4 w_i^2 \right)^{1/2} = \|\mathbf{v}\|_2 \|\mathbf{w}\|_2$ as required.

Problem 3

This question will introduce you to three different ways of thinking about matrix multiplication. For all three parts, let $U \in \mathbb{R}^{m\times n}$, $V \in \mathbb{R}^{m\times r}$ and $W \in \mathbb{R}^{r\times n}$. By $\mathbf{u}_i$ (respectively $\mathbf{v}_i$ and $\mathbf{w}_i$) I mean the i-th column of U (respectively V and W).

1. Show that $U^\top V = \begin{bmatrix} \mathbf{u}_1^\top \mathbf{v}_1 & \mathbf{u}_1^\top \mathbf{v}_2 & \dots & \mathbf{u}_1^\top \mathbf{v}_r \\ \vdots & \ddots & \ddots & \vdots \\ \mathbf{u}_n^\top \mathbf{v}_1 & \mathbf{u}_n^\top \mathbf{v}_2 & \dots & \mathbf{u}_n^\top \mathbf{v}_r \end{bmatrix}$ *(Recall that $\mathbf{u}_i^\top \mathbf{v}_j$ is the dot product between $\mathbf{u}_i$ and $\mathbf{v}_j$)*
2. Show that $UW^\top = \sum_{i=1}^{n} \mathbf{u}_i \mathbf{w}_i^\top$.
3. Show that $U^\top V = [U^\top \mathbf{v}_1, U^\top \mathbf{v}_2, \dots, U^\top \mathbf{v}_r]$.

Solution:

1. Let Z denote the matrix in the RHS. By definition, we have that $(U^\top V)_{ij} = \sum_{l=1}^{m} U_{il}^\top V_{lj}$. But $U_{il}^\top$ for $1 \le l \le m$ is the i-th row of $U^\top$ i.e. it is the i-th column of U. But then we can write $\sum_{l=1}^{m} U_{il}^\top V_{lj} = \langle \mathbf{u}_i, \mathbf{v}_j \rangle = \mathbf{u}_i^\top \mathbf{v}_j = Z_{ij}$ for $1 \le i \le n$ and $1 \le j \le r$.
2. We consider the i,j-th element of the RHS, $\left(\sum_{l=1}^{n} \mathbf{u}_l \mathbf{w}_l^\top\right)_{ij} = \sum_{l=1}^{n} (\mathbf{u}_l \mathbf{w}_l^\top)_{ij} = \sum_{l=1}^{n} \mathbf{u}_{l_i} \mathbf{w}_{l_j} = \sum_{l=1}^{n} U_{il} W_{jl} = \sum_{l=1}^{n} U_{il} W_{lj}^\top = (UW^\top)_{ij}$ by definition. $\therefore UW^\top = \sum_{i=1}^{n} \mathbf{u}_i \mathbf{w}_i^\top$.
3. Similarly, we have $\big([U^\top \mathbf{v}_1, U^\top \mathbf{v}_2, \dots, U^\top \mathbf{v}_r]\big)_{ij} = (U^\top \mathbf{v}_j)_i = \sum_{l=1}^{n} U_{il}^\top \mathbf{v}_{j_l} = \sum_{l=1}^{n} U_{il}^\top V_{lj} = (U^\top V)_{ij}$ for every $1 \le i \le n$ and $1 \le j \le r$.

Problem 4

We say that a function $d : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}_+$ is a metric if:

1. $d(\mathbf{u}, \mathbf{v}) = d(\mathbf{v}, \mathbf{u})$ for all $\mathbf{u}, \mathbf{v} \in \mathbb{R}^n$. *(symmetry)*
2. $d(\mathbf{u}, \mathbf{v}) = 0$ if and only if $\mathbf{u} = \mathbf{v}$. *(definiteness)*
3. $d(\mathbf{u}, \mathbf{w}) \leq d(\mathbf{u}, \mathbf{v}) + d(\mathbf{v}, \mathbf{w})$ for all $\mathbf{u}, \mathbf{v}, \mathbf{w} \in \mathbb{R}^n$ *(triangle inequality).*

Show that if $\|\cdot\|$ is a norm on $\mathbb{R}^n$, then $d(\cdot, \cdot)$ defined by $d(\mathbf{u}, \mathbf{v}) = \|\mathbf{u} - \mathbf{v}\|$ is a metric on $\mathbb{R}^n$.

Solution: Since $(\mathbf{u} - \mathbf{v}) = -1 \times (\mathbf{v} - \mathbf{u})$, by the *positive homogeneity* of norms we have that $d(\mathbf{u}, \mathbf{v}) = \|\mathbf{u} - \mathbf{v}\| = |-1| \times \|\mathbf{v} - \mathbf{u}\| = d(\mathbf{v}, \mathbf{u})$ i.e. d is *symmetric.* Moreover, the *definiteness* of norms means that we must have $\mathbf{u} - \mathbf{v} = 0$ when $d(\mathbf{u}, \mathbf{v}) = 0$, so d is *definite* as well.

Finally, $d(\mathbf{u}, \mathbf{v}) + d(\mathbf{v}, \mathbf{w}) = \|\mathbf{u} - \mathbf{v}\| + \|\mathbf{v} - \mathbf{w}\| \geq \|\mathbf{u} - \mathbf{v} + \mathbf{v} - \mathbf{w}\| = \|\mathbf{u} - \mathbf{w}\| = d(\mathbf{u}, \mathbf{w})$ by the *triangle inequality* of norms. $\therefore d(\mathbf{u}, \mathbf{v}) = \|\mathbf{u} - \mathbf{v}\|$ is a metric.

Problem 5

Prove that $\lim_{p\to\infty} \|\mathbf{v}\|_p = \|\mathbf{v}\|_\infty$. Recall that $\|\cdot\|_\infty$ is the max-norm defined as $\|\mathbf{v}\|_\infty := \max_i |v_i|$.

Solution:

$$\begin{aligned}\|\mathbf{v}\|_p &= \left(\sum_{i=1}^{n} |v_i|^p\right)^{1/p} \qquad \text{Let } |v_{i^*}| = \max_i |v_i| \\ &= \left(|v_{i^*}|^p \left(1 + \sum_{i\neq i^*} \frac{|v_i|^p}{|v_{i^*}|^p}\right)\right)^{1/p}\end{aligned}$$

Note that $0 \leq \sum_{i\neq i^*} \frac{|v_i|^p}{|v_{i^*}|^p} \leq \sum_{i\neq i^*} \frac{|v_{i^*}|^p}{|v_{i^*}|^p} = n-1$, which gives us

$$(|v_{i^*}|^p(1+0))^{1/p} \leq \left(|v_{i^*}|^p \left(1 + \sum_{i\neq i^*} \frac{|v_i|^p}{|v_{i^*}|^p}\right)\right)^{1/p} \leq (|v_{i^*}|^p(1+n-1))^{1/p}$$

$$\implies |v_{i^*}| \leq \left(|v_{i^*}|^p \left(1 + \sum_{i\neq i^*} \frac{|v_i|^p}{|v_{i^*}|^p}\right)\right)^{1/p} \leq |v_{i^*}| \cdot n^{1/p}$$

Taking the limit as $p \to \infty$, we observe that $n^{1/p} \to 1$, and thus using the squeeze theorem, we can conclude that $\lim_{p\to\infty} \|\mathbf{v}\|_p = \|\mathbf{v}\|_\infty$ (where $\|\mathbf{v}\|_\infty = \max_i |v_i|$).

Problem 6

Suppose that:

$$A = LU \text{ where } L = \begin{bmatrix} 1 & 0 & 0 \\ \ell_{21} & 1 & 0 \\ \ell_{31} & \ell_{32} & 1 \end{bmatrix} \text{ and } U = \begin{bmatrix} u_{11} & u_{12} & u_{13} \\ 0 & u_{22} & u_{23} \\ 0 & 0 & u_{33} \end{bmatrix}$$

Find an easy formula for the determinant $\det(A)$, in terms of the components of L and U.

Solution: We know that $\det(A) = \det(LU) = \det(L)\det(U)$. Now, $\det(L) = 1(1\cdot 1 - \ell_{32}\cdot 0) - 0(\cdots) + 0(\cdots) = 1$ and $\det(U) = u_{11}(u_{22}\cdot u_{33} - u_{23}\cdot 0) - u_{12}(0\cdot u_{33} - u_{23}\cdot 0) + u_{13}(0\cdot u_{22} - 0\cdot 0) = u_{11}u_{22}u_{33}$. Hence, we get $\det(A) = u_{11}u_{22}u_{33}$.

Problem 7

In class we defined the condition number of an invertible square matrix as $\kappa(A) = \|A\|_2\|A^{-1}\|_2$. Assume in addition that A is symmetric, so that A is unitarily diagonalizable. In this question you'll show that $\kappa(A) = \frac{\lambda_{\max}(A)}{\lambda_{\min}(A)}$, where $\lambda_{\max}(\cdot)$ (resp. $\lambda_{\min}(\cdot)$) denotes the largest-in-magnitude (resp. smallest-in-magnitude) eigenvalue.

1. Show that $\|A\|_2 = \lambda_{\max}(A)$. *(Don't overthink this, it's essentially "Special Case 1" from slide 8 of Lecture 2).*
2. Show that if $\lambda_1, \ldots, \lambda_n$ are the eigenvalues of A then $\lambda_1^{-1}, \ldots, \lambda_n^{-1}$ are the eigenvalues of A^{-1}.
3. Conclude that $\kappa(A) = \frac{\lambda_{\max}(A)}{\lambda_{\min}(A)}$

Solution:

1. We know that the 2-norm is the special case of induced p-norms where $\|A\|_2 = \lambda_{\max}(A^\top A)^{1/2}$. Since A is symmetric i.e. $A = A^\top$ we have that $\|A\|_2 = \lambda_{\max}(A^2)^{1/2}$. Moreover, note that $A^2\mathbf{v} = A(A\mathbf{v}) = A\lambda_{\max}(A)\mathbf{v} = \lambda_{\max}(A)(A\mathbf{v}) = \lambda_{\max}(A)^2\mathbf{v}$ where $\mathbf{v}$ is the associated eigenvector for $\lambda_{\max}(A)$. In particular, we have that $\lambda_{\max}(A^2) = \lambda_{\max}(A)^2$ and thus $\|A\|_2 = \lambda_{\max}(A)^{2\times 1/2} = \lambda_{\max}(A)$.
2. Let λ be an arbitrary eigenvalue of A and let $\mathbf{v}$ be the associated non-zero eigenvector. Then we have $A\mathbf{v} = \lambda\mathbf{v}$. Multiplying both sides by $A^{-1}\lambda^{-1}$, note that we get $\lambda^{-1}\mathbf{v} = A^{-1}\mathbf{v}$ i.e. λ^{-1} is an eigenvalue of A^{-1}. $\therefore$ WLOG if $\lambda_1, \ldots, \lambda_n$ are the eigenvalues of A then $\lambda_1^{-1}, \ldots, \lambda_n^{-1}$ are the eigenvalues of A^{-1}.
3. From above we have $\|A\|_2 = \lambda_{\max}(A)$ and that $\|A^{-1}\|_2 = \lambda_{\max}(A^{-1})$. In addition, using part 2 we deduce that $\lambda_{\max}(A^{-1}) = \lambda_{\min}(A)^{-1}$ and thus get $\|A^{-1}\|_2 = \lambda_{\min}(A)^{-1}$. Then by definition, $\kappa(A) = \|A\|_2\|A^{-1}\|_2 = \frac{\lambda_{\max}(A)}{\lambda_{\min}(A)}$.

Problem 8

Consider the problem $A\mathbf{x} = \mathbf{b}$ where:

$$A = \begin{bmatrix} 3 & 1 & 0 \\ 1 & 4 & 0 \\ 0 & 0 & \epsilon \end{bmatrix} \quad \text{and} \quad \mathbf{b} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$$

Compute the condition number of A, $\kappa(A)$. *Do this by using the results of Question 7.*

Solution: We first find the eigenvalues of A below. We want

$$\begin{aligned} \det(A - \lambda I) = 0 \implies \det\left(\begin{bmatrix} 3-\lambda & 1 & 0 \\ 1 & 4-\lambda & 0 \\ 0 & 0 & \epsilon-\lambda \end{bmatrix}\right) &= 0 \\ \implies (3-\lambda)(4-\lambda)(\epsilon-\lambda) - (\epsilon-\lambda) &= 0 \\ \implies (\epsilon-\lambda)[(3-\lambda)(4-\lambda) - 1] &= 0 \\ \implies (\epsilon-\lambda)(11 - 7\lambda + \lambda^2) &= 0 \end{aligned}$$

$\therefore$ A has eigenvalues $\lambda_1 = \epsilon, \lambda_2 \approx 2.382$ and $\lambda_3 \approx 4.618$. ϵ is typically small and if this is the case then it largely determines the condition number as $\kappa(A) = \frac{4.618}{\epsilon}$. For completeness, when $2.382 \leq |\epsilon| < 4.618$ we have $\kappa(A) = \frac{4.618}{2.382}$, and when $|\epsilon| \geq 4.618, \kappa(A) = \frac{\epsilon}{3.618}$.

Problem 9

Consider the least squares problem $\mathbf{x}^* = \arg\min_{\mathbf{x}\in\mathbb{R}^4} \|A\mathbf{x} - \mathbf{b}\|_2$ where A and $\mathbf{b}$ are given below.

$$A = \begin{bmatrix} 0.7922 & 0.6787 & 0.7060 & 0.6948 \\ 0.9595 & 0.7577 & 0.0318 & 0.3171 \\ 0.6557 & 0.7431 & 0.2769 & 0.9502 \\ 0.0357 & 0.3922 & 0.0462 & 0.0344 \\ 0.8491 & 0.6555 & 0.0971 & 0.4387 \\ 0.9340 & 0.1712 & 0.8235 & 0.3816 \end{bmatrix} \qquad \mathbf{b} = \begin{bmatrix} 0.7655 \\ 0.7952 \\ 0.1869 \\ 0.4898 \\ 0.4456 \\ 0.6463 \end{bmatrix}$$

$$Q = \begin{bmatrix} -0.4191 & -0.1861 & -0.5979 & -0.1311 & -0.2584 & -0.5901 \\ -0.5076 & -0.1321 & 0.5495 & -0.3017 & -0.5548 & 0.1553 \\ -0.3469 & -0.4172 & -0.1468 & 0.7373 & -0.0250 & 0.3739 \\ -0.0189 & -0.5331 & -0.2732 & -0.5822 & 0.3160 & 0.4494 \\ -0.4492 & -0.0950 & 0.3924 & 0.0215 & 0.7089 & -0.3637 \\ -0.4941 & 0.6933 & -0.3005 & -0.0938 & 0.1500 & 0.3919 \end{bmatrix}$$

$$R_1 = \begin{bmatrix} -1.8902 & -1.3134 & -0.8595 & -1.1681 \\ 0 & -0.6892 & 0.2859 & -0.3632 \\ 0 & 0 & -0.6673 & -0.3327 \\ 0 & 0 & 0 & 0.4674 \end{bmatrix}$$

1. Find $\mathbf{x}^*$ using the QR decomposition method (Use the Q and R given above, which are such that $A = QR$ where $R = \begin{bmatrix} R_1 \\ 0 \end{bmatrix}$). Show your work.

2. Check your work by using matlab or python to solve the least squares problem.

Solution:

1. We know that solving the least squares problem $\mathbf{x}^* = \arg\min_{\mathbf{x}\in\mathbb{R}^4} \|A\mathbf{x} - \mathbf{b}\|_2$ is equivalent to solving the easier problem $\mathbf{x}^* = \arg\min_{\mathbf{x}\in\mathbb{R}^4} \|R\mathbf{x} - Q^\top\mathbf{b}\|_2^2$. We first compute

$$Q^\top\mathbf{b} = \begin{bmatrix} \mathbf{b}_1 \\ \mathbf{b}_2 \end{bmatrix} = \begin{bmatrix} -1.318 \\ -0.181 \\ -0.201 \\ -0.539 \\ -0.076 \\ 0.053 \end{bmatrix} \text{ where } \mathbf{b}_1 = \begin{bmatrix} -1.318 \\ -0.181 \\ -0.201 \\ -0.539 \end{bmatrix}.$$

To minimize $\arg\min_{\mathbf{x}\in\mathbb{R}^4} \|R\mathbf{x} - Q^\top\mathbf{b}\|_2^2 = \|R_1\mathbf{x} - \mathbf{b}_1\|_2^2 + \|\mathbf{b}_2\|_2^2$, we solve $R_1\mathbf{x} = \mathbf{b}_1$ to find $\mathbf{x}^*$. So we have

$$\begin{bmatrix} -1.8902 & -1.3134 & -0.8595 & -1.1681 \\ 0 & -0.6892 & 0.2859 & -0.3632 \\ 0 & 0 & -0.6673 & -0.3327 \\ 0 & 0 & 0 & 0.4674 \end{bmatrix} \times \mathbf{x}^* = \begin{bmatrix} -1.318 \\ -0.181 \\ -0.201 \\ -0.539 \end{bmatrix}.$$

$$\therefore \mathbf{x}^* = \begin{bmatrix} 0.154 \\ 1.233 \\ 0.876 \\ -1.152 \end{bmatrix}$$

2. Solving with numpy's least squares algorithm gives the same result $\mathbf{x}^* = \begin{bmatrix} 0.154 \\ 1.233 \\ 0.876 \\ -1.152 \end{bmatrix}$.

The code used for the above is reproduced below.

```
import numpy as np

A = np.genfromtxt('data/A.csv', delimiter=',')
Q = np.genfromtxt('data/Q.csv', delimiter=',')
b = np.genfromtxt('data/b.csv', delimiter=',')
R1 = np.genfromtxt('data/R1.csv', delimiter=',')

print(latexify(Q.T @ b))

b1 = (Q.T @ b)[:4]
x_star = np.linalg.solve(R1, b1)
print(latexify(x_star))

# checking results using least squares solver
x_lsq = np.linalg.lstsq(A, b)[0]
print(latexify(x_lsq))
```

Problem 10

Let $A \in \mathbb{R}^{m\times n}$. Show that $\|A\|_F = \sqrt{\sum_{i=1}^n \sigma_i(A)^2}$. You will need to use the fact that $\|\cdot\|_F$ is unitarily invariant: $\|UA\|_F = \|A\|_F$ and $\|AV\|_F = \|A\|_F$ for all unitary matrices $U \in \mathbb{R}^{m\times m}$ and $V \in \mathbb{R}^{n\times n}$.

Solution: We know that $\|A\|_F^2 = \text{trace}(A^\top A)$. Using the Singular Value Decomposition $A = U\Sigma V^\top$, we get $\|A\|_F^2 = \text{trace}(V\Sigma^\top U^\top U\Sigma V^\top) = \text{trace}(V\Sigma^\top \Sigma V^\top) = \text{trace}(V^\top V\Sigma^\top\Sigma) = \text{trace}(\Sigma^\top\Sigma)$. Since $\text{trace}(\Sigma^\top\Sigma) = \displaystyle\sum_{i=1}^n \sigma_i(A)^2$, we get $\|A\|_F = \sqrt{\sum_{i=1}^n \sigma_i(A)^2}$ as wanted.

Problem 11

Recall that $A^{(k)} = \sum_{i=1}^{k} \sigma_i \mathbf{u}_i \mathbf{v}_i^\top$ is the best rank k approximation to A. Express the following quantities in terms of the singular values of A:

1. $\|A_k\|_F$.
2. $\|A_k\|_2$.
3. $\|A - A_k\|_F$.
4. $\|A - A_k\|_2$.

Solution: Let $\sigma_1 \geq \sigma_2 \geq ... \geq \sigma_r$.

1. $\|A_k\|_F = \| \sum_{i=1}^{k} \sigma_i \mathbf{u}_i \mathbf{v}_i^\top \|_F = \sqrt{\sum_{i=1}^{k} \sigma_i(A)^2}$
2. $\|A_k\|_2 = \| \sum_{i=1}^{k} \sigma_i \mathbf{u}_i \mathbf{v}_i^\top \|_2 = \sigma_1(A)$
3. $\|A - A_k\|_F = \| \sum_{i=k+1}^{r} \sigma_i \mathbf{u}_i \mathbf{v}_i^\top \|_F = \sqrt{\sum_{i=k+1}^{r} \sigma_i(A)^2}$
4. $\|A - A_k\|_2 = \| \sum_{i=k+1}^{r} \sigma_i \mathbf{u}_i \mathbf{v}_i^\top \|_2 = \sigma_{k+1}(A)$

Problem 12

Compute the SVD of $A = \begin{bmatrix} 1 & 0 \\ 2 & 7 \\ 3 & 5 \end{bmatrix}$ by computing the eigenvalues and eigenvectors of $A^\top A$. Do this by hand and show your work.

Solution: $A^\top A = \begin{bmatrix} 14 & 29 \\ 29 & 74 \end{bmatrix}$ so $\det(A^\top A - \lambda I) = \det\left(\begin{bmatrix} 14-\lambda & 29 \\ 29 & 74-\lambda \end{bmatrix}\right) = 0$.
$\therefore\ (14-\lambda)(74-\lambda) - 841 = 0 \implies \lambda^2 - 88\lambda + 195 = 0$ so that we have $\lambda_1 = 45 + \sqrt{1741} \approx 85.725$ and $\lambda_2 = 45 - \sqrt{1741} \approx 2.275$. The associated eigenvectors $\mathbf{v}_1$ and $\mathbf{v}_2$ such that $A^\top A\mathbf{v}_1 = \lambda_1\mathbf{v}_1$ and $A^\top A\mathbf{v}_2 = \lambda_2\mathbf{v}_2$ then are $\mathbf{v}_1 = (-0.375, -0.927)^\top$ and $\mathbf{v}_2 = (-0.927, 0.375)^\top$. Moreover, $\sigma_1 = \sqrt{\lambda_1} = 9.259$ and $\sigma_2 = \sqrt{\lambda_2} = 1.508$.
Finally, we use $A\mathbf{v} = \sigma\mathbf{u}$ to find the right singular vectors resulting in the following decomposition of $A = U\Sigma V^\top$:

$$U = \begin{bmatrix} -0.04 & 0.615 & -0.788 \\ -0.782 & -0.51 & -0.358 \\ -0.622 & 0.601 & 0.501 \end{bmatrix}, \Sigma = \begin{bmatrix} 9.259 & 0 \\ 0 & 1.508 \end{bmatrix} \text{ and } V^\top = \begin{bmatrix} -0.375 & -0.927 \\ 0.927 & -0.375 \end{bmatrix}.$$

Problem 13

In class we used PCA to analyze the "iris" data set. In this question you will repeat the analysis for another classic machine learning data set, "wine". Instead of

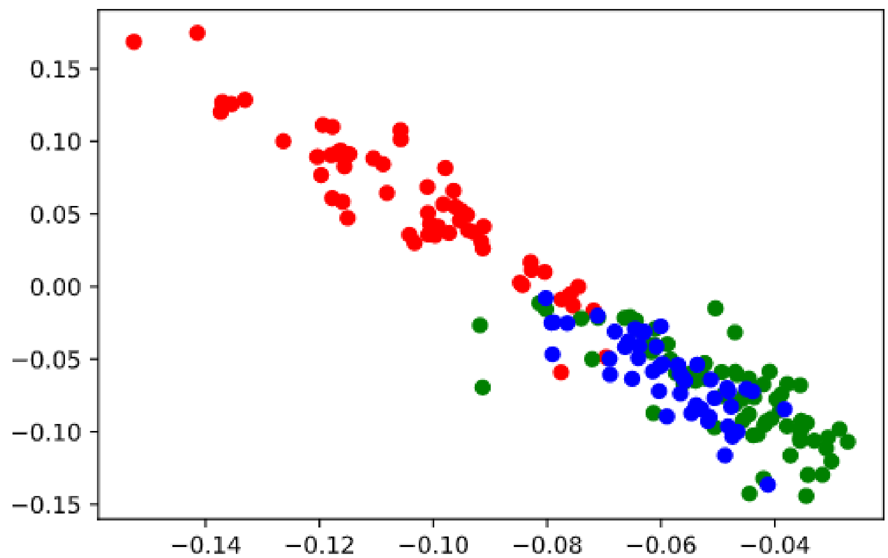
```
iris = load_iris()
```

use the command:

```
wine = load_wine()
```

in the Jupyter notebook from class. Make appropriate adjustments to the code from class, and attach a plot of the loadings on to the first two singular vectors of this data set. Color the data points to indicate which of the three classes each data point belongs to.

Solution: The plot below is on the loadings of the leading two singular vectors of the wine dataset, colored by the classes the wines below to. The code used to create this figure is presented below.




```
import numpy as np
import matplotlib.pyplot as plt
from sklearn.datasets import load_wine

wine = load_wine()
X = wine.data
y = wine.target
mu = np.mean(X, axis=1, keepdims=True)
X_temp = X - mu
X_tilde = X_temp/np.max(X_temp,axis = 0)
U,S,Vh = np.linalg.svd(X_tilde)
Loading_on_sing_vec_1 = U[:,0]
Loading_on_sing_vec_2 = U[:,1]
colors = ['red', 'green', 'blue']
for i in range(0,3):
    plt.scatter(Loading_on_sing_vec_1[y == i],Loading_on_sing_vec_2[y == i],color = colors[i])
plt.show()
```