Homework 13

Self grades are due at 11 PM on May 4, 2023.

1. **Project Logistics**

   Fill out this form to let us know whether:

   - you plan on doing the project in a group;
   - you plan on doing the project by yourself; or
   - you are not planning to do the project;

   as well as some auxiliary information. Even if you do not plan to do the project, you **must** fill out the form.

   To get credit for the problem, attach a screenshot of the filled-out form to the PDF you submit to Gradescope.

2. Wasserstein distance between distributions

The Wasserstein distance is a measure of distance between probability distributions. The Wasserstein distance can roughly be thought of as the cost of turning one distribution to another distribution by moving probability mass around from one location to another. It is also sometimes called the earth-mover distance, because it may be visualized as the cost of moving a pile of dirt from one configuration to another.

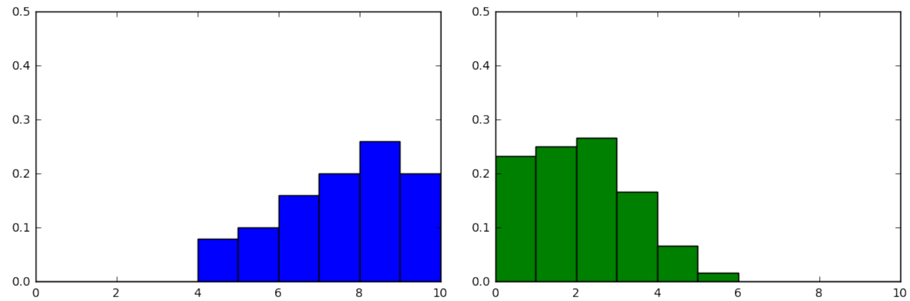


Figure 1: Visualization of μ histogram on left and ν histogram on right.

Let $n \in \mathbb{N}$. We define two discrete probability distributions $\vec{\mu} = (\mu_1, \cdots, \mu_n)$ and $\vec{\nu} = (\nu_1, \cdots, \nu_n)$; that is, $\mu_i, \nu_i \geq 0$ and $\sum_i \mu_i = \sum_i \nu_i = 1$.

We define $C \in \mathbb{R}^{n \times n}$ to be a cost matrix where $c_{ij} \geq 0$ is the cost of transporting one unit of probability mass from location $i \in \{1, \cdots, n\}$ to location $j \in \{1, \cdots, n\}$. We define a matrix $M \in \mathbb{R}^{n \times n}$ where $m_{ij} \geq 0$ denotes the quantity of probability mass to be moved from location i to location j. In summary, if we move m_{ij} units of probability mass from location i to location j, we incur cost $c_{ij} m_{ij}$.

In addition, the M matrix satisfies the following conditions. Row i of M indicates where all the probability mass in location i in the $\vec{\mu}$ distribution ends up. Hence, the sum of all the entries in row i must equal μ_i. Similarly, column j indicates where all the probability mass in location j in the $\vec{\nu}$ distribution came from. Hence, the sum of all the entries in column j must equal ν_j. We can summarize these conditions in math:

$$M\vec{1} = \vec{\mu} \tag{1}$$

$$M^\top \vec{1} = \vec{\nu}, \tag{2}$$

where $\vec{1}$ is a vector of 1s.

(a) What is the total cost of transporting the mass $\vec{\mu}$ into $\vec{\nu}$ by following the transportation plan dictated by the matrix M?

Solution: The total cost by following the transportation plan is $\sum_{i,j} c_{ij} m_{ij} = \text{Tr}(C^T M) = \langle C, M \rangle_F$ where $\langle C, M \rangle_F$ is the Frobenius inner product.

(b) Given the cost matrix C, write the optimization problem of finding the transportation plan $M^\star$ with minimal total cost. What type of optimization problem is it? (LP, QP, $\cdots$?).

Solution: To find $M^\star$ that incurs minimal total cost given the fixed cost matrix C, we can formulate a minimization problem over M. The objective of the problem is the total cost we derived in part (a), $\langle C, M \rangle_F$. This minimization problem is also subject to the following constraints from the definition of M:

$$M\vec{1} = \mu$$

$$M^\top \vec{1} = \nu$$
$$M \geq 0$$

So, finially we piece everything together and write the minimization problem as:

$$\begin{aligned} M^\star = \operatorname*{argmin}_{M} \quad & \langle C, M \rangle_F \\ \text{s.t.} \quad & M\vec{1} = \mu \\ & M^\top \vec{1} = \nu \\ & M \geq 0 \end{aligned}$$

Since both the objective and the constraints are affine in M, this optimization problem is an LP.

Now, we apply the idea of Wasserstein distance to document similarity as illustrated in Fig. 2. Here, our application is that we want to identify words in two different documents that are most similar. This is mostly just a fun application, but may be of interest if you are trying to compare documents that are identical but in different languages. Here we consider a contrived example.

Natural Language Processing techniques have standard tools for converting words into vectors and embedding them in vector spaces, so that we can use machine learning and optimization tools on them. One such embedding is called *word2vec*. Assume we are provided with a *word2vec* embedding for the words in two documents. The word travel cost c_{ij} between word i and word j is the Euclidean distance $\|x_i - x_j\|_2$ in the word embedding space. We can compute the similarity between two documents as the minimum cumulative cost required to move all non-stop words from one document to the other.

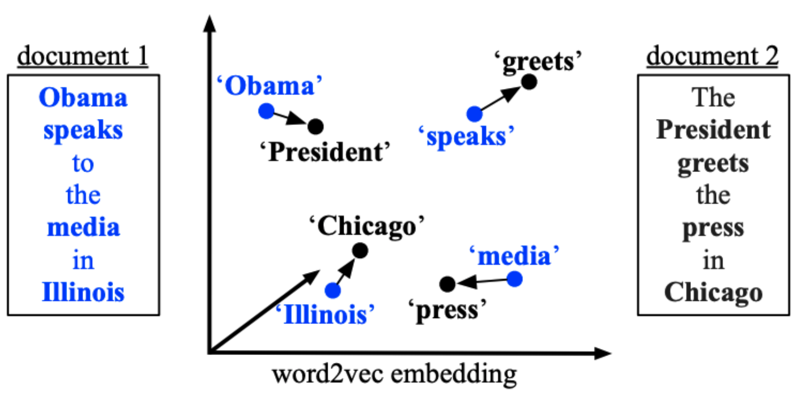
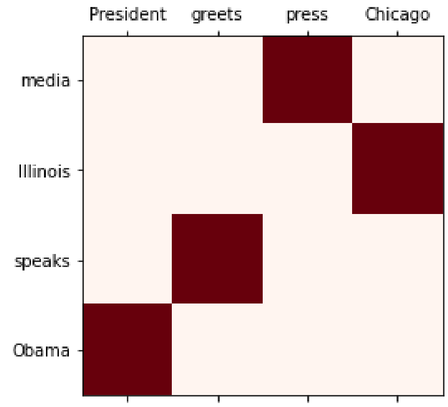


Figure 2: An illustration of the Wasserstein distance. All non-stop words (**bold**) of both documents are embedded into a *word embedding* space. The similarity between the two documents is the minimum cumulative distance that all words in document 1 need to travel to exactly match document 2.

(c) Using the `text_kantorovich.ipynb` Juypter notebook, implement the calculation of the Wasserstein distance in the notebook and use the provided code to visualize the resulting matrix M. Comment on the results.

Solution: Formulating and solving the problem with cvxpy produces the following plot. We see that the matrix M agrees with our intuition and "moves" words to semantically similar words.


President
greets
press
Chicago
media
Illinois
speaks
Obama

3. Linear Quadratic Regulator

In this question, we will derive the Riccati equation for the LQR model studied in class. We first recall the statement of the LQR problem:

$$\begin{aligned}\min_{\vec{x}_t,\vec{u}_t} \quad & \sum_{t=0}^{T-1} \frac{1}{2}\left(\vec{x}_t^\top Q \vec{x}_t + \vec{u}_t^\top R \vec{u}_t\right) + \frac{1}{2}\vec{x}_T^\top Q \vec{x}_T && (3)\\ \text{s.t.} \quad & \vec{x}_{t+1} = A\vec{x}_t + B\vec{u}_t && (4)\\ & \vec{x}_0 = \vec{x}_{\text{init}} && (5)\end{aligned}$$

where $\vec{x}_t$ is thought of as the state of the system and $\vec{u}_t$ is the control input at time t and the matrices A and B define the dynamics of the system. Here $Q, R \succ 0$ are symmetric positive definite matrices determining how the state and control affect the cost. While the problem can be solved as a quadratic program, we will now take a slightly different approach. We start by defining the functions, J_k for $0 \leq k \leq T$, as follows:

$$\begin{aligned}J_k(\vec{x}) = \min_{\{\vec{u}_t\}_{t=k}^{T-1}} \quad & \sum_{t=k}^{T-1} \frac{1}{2}\left(\vec{x}_t^\top Q \vec{x}_t + \vec{u}_t^\top R \vec{u}_t\right) + \frac{1}{2}\vec{x}_T^\top Q \vec{x}_T && (6)\\ \text{s.t.} \quad & \vec{x}_{t+1} = A\vec{x}_t + B\vec{u}_t && (7)\\ & \vec{x}_k = \vec{x}. && (8)\end{aligned}$$

J_k can be thought of as the minimum cost that we would incur from time k assuming that we start at state $\vec{x}_k = \vec{x}$. We can now decompose J_k for $0 \leq k \leq T-1$ further as follows:

$$\begin{aligned}J_k(\vec{x}) = \min_{\vec{u}_k} \quad & \frac{1}{2}\left(\vec{x}_k^\top Q \vec{x}_k + \vec{u}_k^\top R \vec{u}_k\right) + \min_{\{\vec{u}_t\}_{t=k+1}^{T-1}} \sum_{t=k+1}^{T-1} \frac{1}{2}\left(\vec{x}_t^\top Q \vec{x}_t + \vec{u}_t^\top R \vec{u}_t\right) + \frac{1}{2}\vec{x}_T^\top Q \vec{x}_T && (9)\\ \text{s.t.} \quad & \vec{x}_{t+1} = A\vec{x}_t + B\vec{u}_t && (10)\\ & \vec{x}_k = \vec{x}. && (11)\end{aligned}$$

Note that in particular, the first constraint implies that $\vec{x}_{k+1} = A\vec{x}_k + B\vec{u}_k$. Therefore, the above characterization gives the following decomposition:

$$J_k(\vec{x}) = \min_{\vec{u}} \frac{1}{2}\left(\vec{x}^\top Q \vec{x} + \vec{u}^\top R \vec{u}\right) + J_{k+1}(A\vec{x} + B\vec{u}). \tag{12}$$

This equation is called the *Bellman equation* and is solved by dynamic programming, as we will show in the problem.

We will see that the functions, J_k, are all in fact quadratic functions in $\vec{x}$ and this will give us convenient ways to derive the optimal control inputs at each time.

(a) First, we will show by reverse induction that each of the functions J_k for $0 \leq k \leq T$ are convex quadratics. In particular, prove that $J_k(\vec{x}) = \frac{1}{2}\vec{x}^\top Q_k \vec{x}$ for some $Q_k \succ 0$ and determine the value of Q_k in terms of Q_{k+1}.

HINT: $J_T(\vec{x}) = \frac{1}{2}\vec{x}^\top Q \vec{x}$. Therefore, $Q_T = Q$. Also can use (12) above, and substitute $J_{k+1}(A\vec{x} + B\vec{u}) = \frac{1}{2}(A\vec{x} + B\vec{u})^\top Q_{k+1}(A\vec{x} + B\vec{u})$. Then solve the resulting QP to find the optimal $\vec{u}$.

HINT: You should get the following recursion for Q_k:

$$Q_k = Q + A^\top Q_{k+1} A - A^\top Q_{k+1} B (R + B^\top Q_{k+1} B)^{-1} B^\top Q_{k+1} A. \tag{13}$$

and you may assume that $A^\top Q_{k+1}A - A^\top Q_{k+1}B(R+B^\top Q_{k+1}B)^{-1}B^\top Q_{k+1}A$ is positive semi-definite when you try to establish that Q_k is positive definite.

Solution: We will prove the statement by reverse induction starting with $k = T$ and proving that J_k is a quadratic in the given form for $k = l$ given that it is true for $k = l + 1$.

Note that the statement is true for $k = T$ as $J_T(\vec{x}) = \frac{1}{2}\vec{x}^\top Q\vec{x}$. Now, assume that the statement is true for $k = l + 1$, we will show that it holds for $k = l$. Observe that we can use our recursive definition of J_k to obtain:

$$\begin{aligned}
J_l(\vec{x}) &= \min_{\vec{u}} \frac{1}{2}\left(\vec{x}^\top Q\vec{x} + \vec{u}^\top R\vec{u}\right) + J_{l+1}(A\vec{x} + B\vec{u}) && (14)\\
&= \min_{\vec{u}} \frac{1}{2}\left(\vec{x}^\top Q\vec{x} + \vec{u}^\top R\vec{u} + (A\vec{x} + B\vec{u})^\top Q_{l+1}(A\vec{x} + B\vec{u})^\top\right) && (15)\\
&= \min_{\vec{u}} \frac{1}{2}\left(\vec{x}^\top Q\vec{x} + \vec{u}^\top R\vec{u} + \vec{x}^\top A^\top Q_{l+1}A\vec{x} + \vec{u}^\top B^\top Q_{l+1}B\vec{u} + 2\vec{u}^\top B^\top Q_{l+1}A\vec{x}\right) && (16)\\
&= \min_{\vec{u}} \frac{1}{2}\left(\vec{x}^\top (Q + A^\top Q_{l+1}A)\vec{x} + \vec{u}^\top (R + B^\top Q_{l+1}B)\vec{u} + 2\vec{u}^\top B^\top Q_{l+1}A\vec{x}\right). && (17)
\end{aligned}$$

Note that since the above equation is a convex quadratic optimization problem in $\vec{u}$ for a fixed $\vec{x}$. Therefore, we can compute its derivative with respect to $\vec{u}$ and set it to 0 to explicitly minimize over $\vec{u}$ to obtain:

$$(R + B^\top Q_{l+1}B)\vec{u}^\star + B^\top Q_{l+1}A\vec{x} = 0 \implies \vec{u}^\star = -(R + B^\top Q_{l+1}B)^{-1}B^\top Q_{l+1}A\vec{x}. \quad (18)$$

By substituting $\vec{u}^\star$ we see that:

$$\begin{aligned}
&(\vec{u}^\star)^\top (R + B^\top Q_{l+1}B)\vec{u}^\star + 2(\vec{u}^\star)^\top B^\top Q_{l+1}A\vec{x} && (19)\\
&= \vec{x}^\top A^\top Q_{l+1}B(R + B^\top Q_{l+1}B)^{-1}(R + B^\top Q_{l+1}B)(R + B^\top Q_{l+1}B)^{-1}B^\top Q_{l+1}A\vec{x} && (20)\\
&\quad - 2\vec{x}^\top A^\top Q_{l+1}B(R + B^\top Q_{l+1}B)^{-1}B^\top Q_{l+1}A\vec{x} && (21)\\
&= -\vec{x}^\top A^\top Q_{l+1}B(R + B^\top Q_{l+1}B)^{-1}B^\top Q_{l+1}A\vec{x}. && (22)
\end{aligned}$$

Therefore, we by substituting this in the above equation, we get the following:

$$\begin{aligned}
J_l(\vec{x}) &= \frac{1}{2}\vec{x}^\top Q_l\vec{x} && (23)\\
Q_l &= Q + A^\top Q_{l+1}A - A^\top Q_{l+1}B(R + B^\top Q_{l+1}B)^{-1}B^\top Q_{l+1}A. && (24)
\end{aligned}$$

The fact that $Q_l \succ 0$ follows from the fact that Q_l is the sum of $Q \succ 0$ and a PSD matrix. We see that the latter part is a PSD matrix as follows: We know that by reverse induction Q_{k+1} is PD and hence has a square root and an inverse as well as an inverse square root denoted by $Q_{k+1}^{-1/2}$. We first have:

$$\begin{aligned}
A^\top Q_{k+1}A - A^\top Q_{k+1}B(R + B^\top Q_{k+1}B)^{-1}B^\top Q_{k+1}A = \\
A^\top Q_{k+1}^{1/2}\left[I - Q_{k+1}^{1/2}B(R + B^\top Q_{k+1}B)^{-1}B^\top Q_{k+1}^{1/2}\right]Q_{k+1}^{1/2}A
\end{aligned}$$

Now we will prove that all the eigenvalues of $Q_{k+1}^{1/2}B(R + B^\top Q_{k+1}B)^{-1}B^\top Q_{k+1}^{1/2}$ are less than or equal to 1 which would establish that the term in the square brackets is PSD. This implies that the whole matrix is PSD as if a matrix P is PSD, then for any matrix L, $L^\top PL$ is also PSD which can easily be checked by evaluating the quadratic form.

Now, as R is PD, it also has an inverse square root.

$$(R + B^\top Q_{k+1}B)^{-1} = (R^{1/2}(I + R^{-1/2}B^\top Q_{k+1}BR^{-1/2})R^{1/2})^{-1}$$

$$= R^{-1/2}(I + R^{-1/2}B^\top Q_{k+1}BR^{-1/2})^{-1}R^{-1/2}$$

where the last step follows as $(ABC)^{-1} = C^{-1}B^{-1}A^{-1}$ Now denote $L = Q_{k+1}^{1/2}BR^{-1/2}$ and let it have an SVD $L = U\Sigma V^\top$ so the term in the middle is

$$\begin{aligned}(I + R^{-1/2}B^\top Q_{k+1}BR^{-1/2})^{-1} &= (I + L^\top L)^{-1} \\ &= (I + V\Sigma^2V^\top)^{-1} \\ &= V(I + \Sigma^2)^{-1}V^\top\end{aligned}$$

Plugging it back in the previous expression we have,

$$\begin{aligned}Q_{k+1}^{1/2}B(R + B^\top Q_{k+1}B)^{-1}B^\top Q_{k+1}^{1/2} &= Q_{k+1}^{1/2}BR^{-1/2}V(I + \Sigma^2)^{-1}R^{-1/2}B^\top Q_{k+1}^{1/2} \\ &= LV(I + \Sigma^2)^{-1}V^\top L^\top \\ &= U\Sigma(I + \Sigma^2)^{-1}\Sigma U^\top\end{aligned}$$

so the eigenvalues of this are $\sigma_i^2/(1 + \sigma_i^2)$ which are clearly all less than or equal to 1. Hence we are done.

(b) Now, show that the expression for Q_l is equivalent for the expression obtained by using the Lagrangian. That is, show that Q_l from the previous part is the same as:

$$Q_l = Q + A^\top(Q_{l+1}^{-1} + BR^{-1}B^\top)^{-1}A. \tag{25}$$

HINT: You may find useful the Sherman-Morrison-Woodbury matrix identity:

$$(M + UWV)^{-1} = M^{-1} - M^{-1}U(W^{-1} + VM^{-1}U)^{-1}VM^{-1}. \tag{26}$$

Solution: We start by manipulating the middle expression in the definition of Q_l found in part (a):

$$\begin{aligned}&A^\top Q_{l+1}A - A^\top Q_{l+1}B(R + B^\top Q_{l+1}B)^{-1}B^\top Q_{l+1}A && (27)\\ &= A^\top(Q_{l+1} - Q_{l+1}B(R + B^\top Q_{l+1}B)^{-1}B^\top Q_{l+1})A && (28)\\ &= A^\top(Q_{l+1}^{-1} + BR^{-1}B^\top)^{-1}A && (29)\end{aligned}$$

where the last equality follows from the Sherman-Morrison-Woodbury identity by substituting $M = Q_{l+1}^{-1}$, $W = R^{-1}$, $U = B$ and $V = B^\top$. By substituting this expression in the definition of Q_l from Part (a), we get the desired conclusion:

$$Q_{l+1} = Q + A^\top(Q_{l+1}^{-1} + BR^{-1}B^\top)^{-1}A. \tag{30}$$

Finally, notice that we can further factorize the above expression as follows:

$$A^\top(Q_{l+1}^{-1} + BR^{-1}B^\top)^{-1}A = A^\top Q_{l+1}^{1/2}(I + Q_{l+1}^{1/2}BR^{-1}B^\top Q_{l+1}^{1/2})^{-1}Q_{l+1}^{1/2}A \tag{31}$$

to obtain

$$Q_{l+1} = Q + A^\top Q_{l+1}^{1/2}(I + Q_{l+1}^{1/2}BR^{-1}B^\top Q_{l+1}^{1/2})^{-1}Q_{l+1}^{1/2}A. \tag{32}$$

4. Soft-Margin SVM

Consider the soft-margin SVM problem,

$$p^\star(C) = \min_{\vec{w}\in\mathbb{R}^m, b\in\mathbb{R}, \vec{\xi}\in\mathbb{R}^n} \quad \frac{1}{2}\|\vec{w}\|_2^2 + C\sum_{i=1}^n \xi_i \tag{33}$$

$$\text{s.t.} \quad 1 - \xi_i - y_i(\vec{x}_i^\top \vec{w} - b) \leq 0, \quad i = 1, 2, \ldots, n \tag{34}$$

$$-\xi_i \leq 0, \quad i = 1, 2, \ldots, n, \tag{35}$$

where $\vec{x}_i \in \mathbb{R}^m$ refers to the i^{th} training data point, $y_i \in \{-1, 1\}$ is its label, and $C \in \mathbb{R}_+$ (i.e. $C > 0$) is a hyperparameter. Let α_i denote the dual variable corresponding to the inequality $1 - \xi_i - y_i(\vec{x}_i^\top \vec{w} - b) \leq 0$ and let β_i denote the dual variable corresponding to the inequality $-\xi_i \leq 0$. The Lagrangian is then given by

$$\mathcal{L}(\vec{w}, b, \vec{\xi}, \vec{\alpha}, \beta) = \frac{1}{2}\|\vec{w}\|_2^2 + C\sum_{i=1}^n \xi_i + \sum_{i=1}^n \alpha_i(1 - \xi_i - y_i(\vec{x}_i^\top \vec{w} - b)) - \sum_{i=1}^n \beta_i \xi_i. \tag{36}$$

Suppose $\vec{w}^\star, b^\star, \vec{\xi}^\star, \vec{\alpha}^\star, \beta^\star$ satisfy the KKT conditions. Classify the following statements as true or false and justify your answers mathematically.

(a) Suppose the optimal solution $\vec{w}^\star, b^\star$ changes when the training point $\vec{x}_i$ is removed. Then originally, we necessarily have $y_i(\vec{x}_i^\top \vec{w}^\star - b^\star) = 1 - \xi_i^\star$.

Solution: True. Since optimal $\vec{w}^\star$ changes if we remove point $\vec{x}_i$ we have $\alpha_i^\star \neq 0$. By complementary slackness we have,

$$\alpha_i^\star(1 - \xi_i^\star - y_i(\vec{x}_i^\top \vec{w}^\star - b^\star)) = 0, \tag{37}$$

which gives,

$$1 - \xi_i^\star - y_i(\vec{x}_i^\top \vec{w}^\star - b^\star) = 0 \tag{38}$$

$$\implies y_i(\vec{x}_i^\top \vec{w}^\star - b^\star) = 1 - \xi_i^\star. \tag{39}$$

(b) Suppose the optimal solution $\vec{w}^\star, b^\star$ changes when the training point $\vec{x}_i$ is removed. Then originally, we necessarily have $\alpha_i^\star > 0$.

Solution: True. Since optimal $\vec{w}^\star$ changes if we remove point $\vec{x}_i$ we have $\alpha_i^\star \neq 0$. Further by dual feasibility we have $\alpha_i^\star \geq 0$ which together gives $\alpha_i^\star > 0$.

(c) Suppose the data points are strictly linearly separable, i.e. there exist $\vec{\widetilde{w}}$ and $\widetilde{b}$ such that for all i,

$$y_i(\vec{x}_i^\top \vec{\widetilde{w}} - \widetilde{b}) > 0. \tag{40}$$

Then $p^\star(C) \to \infty$ as $C \to \infty$.

Solution: False. Since

$$y_i(\vec{x}_i^\top \vec{\widetilde{w}} - \widetilde{b}) > 0, \tag{41}$$

we have for sufficiently small $\epsilon > 0$,

$$y_i(\vec{x}_i^\top \vec{\widetilde{w}} - \widetilde{b}) \geq \epsilon \implies y_i\left(\vec{x}_i^\top \frac{\vec{\widetilde{w}}}{\epsilon} - \frac{\widetilde{b}}{\epsilon}\right) \geq 1. \tag{42}$$

Thus, $\vec{\overline{w}} = \frac{\vec{\widetilde{w}}}{\epsilon}, \overline{b} = \frac{\widetilde{b}}{\epsilon}, \vec{\overline{\xi}} = 0$ is a feasible point with objective value $\frac{1}{2}\left\|\vec{\overline{w}}\right\|_2^2 < \infty$ irrespective of value of C.

5. Ridge Regression Classifier Vs. SVM

In this problem, we explore Ridge Regression as a classifier, and compare it to SVM. Recall Ridge Regression solves the problem

$$\min_{\vec{w}} \|X\vec{w} - \vec{y}\|_2^2 + \lambda \|\vec{w}\|_2^2, \tag{43}$$

where $X \in \mathbb{R}^{m\times n}$, and $\vec{y} \in \mathbb{R}^n$

(a) Ridge Regression as is solves a regression problem. Given data $X \in \mathbb{R}^{m\times n}$ and labels $\vec{y} \in \{-1, 1\}^m$, explain how we might be able to train a Ridge Regression model and use it to classify a test point.

Solution: We have that the optimal $\vec{w}$ from solving ridge regression is given by

$$\vec{w}^* = (X^\top X + \lambda I)^{-1} X^\top \vec{y} \tag{44}$$

Hence given a new data point x_{test}, we look at $x_{\text{test}}^\top \vec{w}^*$ and if it is positive, we say $y_{\text{test}} = 1$ and otherwise we say $y_{\text{test}} = -1$. In other words, we can say $y_{\text{test}} = \operatorname{sign}(x_{\text{test}}^\top \vec{w}^*)$.

(b) Complete the accompanying Jupyter notebook to compare Ridge Regression and SVM.

Solution: See Jupyter notebook for coding solution.

6. Modified SVM

Let $C > 0$. Suppose we have labeled data $(\vec{x}_i, y_i) \in \mathbb{R}^d \times \{-1, 1\}$ for $i = 1, \ldots, n$. For each i, define $\vec{z}_i \doteq y_i \vec{x}_i$. Finally, define $Z \doteq \begin{bmatrix} \vec{z}_1, \ldots, \vec{z}_n \end{bmatrix}^\top \in \mathbb{R}^{n \times d}$.

Recall that the soft-margin support vector machine problem can be expressed using slack variables as

$$\begin{aligned} p_1^\star = \min_{\vec{w}, \vec{s}} \quad & \frac{1}{2} \|\vec{w}\|_2^2 + C \sum_{i=1}^n s_i \\ \text{s.t.} \quad & s_i = \max\{0, 1 - \vec{z}_i^\top \vec{w}\}, \qquad \forall i \in \{1, \ldots, n\}. \end{aligned} \tag{45}$$

In this problem we consider a modified SVM program with a squared penalty:

$$\begin{aligned} p_2^\star = \min_{\vec{w}, \vec{s}} \quad & \frac{1}{2} \|\vec{w}\|_2^2 + \frac{C}{2} \sum_{i=1}^n s_i^2 \\ \text{s.t.} \quad & s_i = \max\{0, 1 - \vec{z}_i^\top \vec{w}\}, \qquad \forall i \in \{1, \ldots, n\}. \end{aligned} \tag{46}$$

We will use another representation of this program, namely one with affine constraints:

$$\begin{aligned} p^\star = \min_{\vec{w}, \vec{s}} \quad & \frac{1}{2} \|\vec{w}\|_2^2 + \frac{C}{2} \|\vec{s}\|_2^2 \\ \text{s.t.} \quad & \vec{s} \geq \vec{0} \\ & \vec{s} \geq \vec{1} - Z\vec{w}, \end{aligned} \tag{47}$$

where the inequality constraints are componentwise (as usual).

(a) Choose the smallest class that problem (47) belongs to (LP/QP/SOCP/etc).

Solution: It is a QP – it has a quadratic objective and affine constraints.

(b) Prove that strong duality holds for (47).

Solution: The objective function is a convex quadratic and the constraints are affine (hence convex) in $\vec{w}$ and $\vec{s}$, so the problem is convex. Furthermore, there is a strictly feasible point – we can construct one by picking any $\vec{w}$ and then picking $\vec{s}$ whose components are large enough to fulfill the inequalities. This is always possible since there is no upper bound on the components of $\vec{s}$. Thus Slater's condition holds, so strong duality holds.

(c) Are the KKT conditions for problem (47) necessary, sufficient or both necessary and sufficient for global optimality?

Solution: The objective function is a convex quadratic and the constraints are affine (hence convex) in $\vec{w}$ and $\vec{s}$, so the problem is convex.

Since the problem is convex, all functions involved are continuously differentiable, and strong duality holds, the KKT conditions are both necessary and sufficient for optimality; that is, they are equivalent to optimality conditions.

(d) Let $\vec{\alpha}$ be the dual variable corresponding to the constraint $\vec{s} \geq \vec{0}$. What is the dimension (i.e., number of entries) of $\vec{\alpha}$?

Solution: $\vec{\alpha} \in \mathbb{R}^n$ since $\vec{s} \in \mathbb{R}^n$.

(e) Show that the Lagrangian $L(\vec{w}, \vec{s}, \vec{\alpha}, \vec{\beta})$ of problem (47), where $\vec{\alpha}$ is the dual variable corresponding to the constraint $\vec{s} \geq \vec{0}$, and $\vec{\beta}$ is the dual variable corresponding to the constraint $\vec{s} \geq \vec{1} - Z\vec{w}$, is equal to

$$L(\vec{w}, \vec{s}, \vec{\alpha}, \vec{\beta}) = \frac{1}{2}\|\vec{w}\|_2^2 + \frac{C}{2}\|\vec{s}\|_2^2 - \vec{s}^\top(\vec{\alpha} + \vec{\beta}) - \vec{w}^\top Z^\top \vec{\beta} + \vec{1}^\top \vec{\beta}. \tag{48}$$

Solution: We have

$$\begin{aligned} L(\vec{w}, \vec{s}, \vec{\alpha}, \vec{\beta}) &= \frac{1}{2}\|\vec{w}\|_2^2 + \frac{C}{2}\|\vec{s}\|_2^2 + \vec{\alpha}^\top(-\vec{s}) + \vec{\beta}^\top(\vec{1} - Z\vec{w} - \vec{s}) & (49)\\ &= \frac{1}{2}\|\vec{w}\|_2^2 + \frac{C}{2}\|\vec{s}\|_2^2 - \vec{s}^\top(\vec{\alpha} + \vec{\beta}) - \vec{w}^\top Z^\top \vec{\beta} + \vec{1}^\top \vec{\beta}. & (50)\end{aligned}$$

(f) Write the KKT conditions for problem (47). Show that if $(\vec{w}^\star, \vec{s}^\star, \vec{\alpha}^\star, \vec{\beta}^\star)$ obey the KKT conditions for problem (47), then

$$\vec{w}^\star = Z^\top \vec{\beta}^\star \qquad \text{and} \qquad \vec{s}^\star = \frac{\vec{\alpha}^\star + \vec{\beta}^\star}{C}. \tag{51}$$

HINT: For the first order/stationarity condition on the Lagrangian you will need to consider partial derivatives with respect to both $\vec{w}$ and $\vec{s}$.

Solution: Let $(\vec{w}^\star, \vec{s}^\star, \vec{\alpha}^\star, \vec{\beta}^\star)$ satisfy the KKT conditions. We have:

- Primal feasibility: $\vec{s}^\star \geq 0$ and $\vec{s}^\star \geq \vec{1} - Z\vec{w}^\star$.
- Dual feasibility: $\vec{\alpha}^\star \geq \vec{0}$, $\vec{\beta}^\star \geq \vec{0}$.
- Complementary slackness: $\alpha_i^\star s_i^\star = 0$ and $\beta_i^\star(1 - \vec{z}_i^\top \vec{w}^\star - s_i^\star) = 0$ for each i.
- Stationarity: $\nabla_{\vec{w}} L(\vec{w}^\star, \vec{s}^\star, \vec{\alpha}^\star, \vec{\beta}^\star) = \vec{0}$ and $\nabla_{\vec{s}} L(\vec{w}^\star, \vec{s}^\star, \vec{\alpha}^\star, \vec{\beta}^\star) = \vec{0}$. These become

$$\vec{0} = \vec{w}^\star - Z^\top \vec{\beta}^\star \tag{52}$$
$$\vec{0} = C\vec{s}^\star - (\vec{\alpha}^\star + \vec{\beta}^\star) \tag{53}$$

which rearrange to the claimed equalities.

(g) Compute the dual function of problem (47) as

$$g(\vec{\alpha}, \vec{\beta}) \doteq L(\vec{w}^\star(\vec{\alpha}, \vec{\beta}), \vec{s}^\star(\vec{\alpha}, \vec{\beta}), \vec{\alpha}, \vec{\beta}) \tag{54}$$

where from the previous part we have that

$$\vec{w}^\star(\vec{\alpha}, \vec{\beta}) = Z^\top \vec{\beta} \qquad \text{and} \qquad \vec{s}^\star(\vec{\alpha}, \vec{\beta}) = \frac{\vec{\alpha} + \vec{\beta}}{C}. \tag{55}$$

Your final expression for $g(\vec{\alpha}, \vec{\beta})$ should not contain any maximizations, minimizations or terms including $\vec{w}$, $\vec{s}$, $\vec{w}^\star$, or $\vec{s}^\star$. It should only contain $\vec{\alpha}$, $\vec{\beta}$, C, Z, and numerical constants.

Solution: The dual function is

$$\begin{aligned} g(\vec{\alpha}, \vec{\beta}) &= L(\vec{w}^\star(\vec{\alpha}, \vec{\beta}), \vec{s}^\star(\vec{\alpha}, \vec{\beta}), \vec{\alpha}, \vec{\beta}) & (56)\\ &= \frac{1}{2}\left\|\vec{w}^\star(\vec{\alpha}, \vec{\beta})\right\|_2^2 + \frac{C}{2}\left\|\vec{s}^\star(\vec{\alpha}, \vec{\beta})\right\|_2^2 - \vec{s}^\star(\vec{\alpha}, \vec{\beta})^\top(\vec{\alpha} + \vec{\beta}) - \vec{w}^\star(\vec{\alpha}, \vec{\beta})^\top Z^\top \vec{\beta} + \vec{1}^\top \vec{\beta} & (57)\\ &= \frac{1}{2}\left\|Z^\top \vec{\beta}\right\|_2^2 + \frac{C}{2}\left\|\frac{\vec{\alpha} + \vec{\beta}}{C}\right\|_2^2 - \left(\frac{\vec{\alpha} + \vec{\beta}}{C}\right)^\top(\vec{\alpha} + \vec{\beta}) - \vec{\beta}^\top Z Z^\top \vec{\beta} + \vec{1}^\top \vec{\beta} & (58)\\ &= -\frac{1}{2}\vec{\beta}^\top Z Z^\top \vec{\beta} - \frac{1}{2C}\left\|\vec{\alpha} + \vec{\beta}\right\|_2^2 + \vec{1}^\top \vec{\beta}. & (59)\end{aligned}$$

(h) Let $\vec{\alpha}^\star$ and $\vec{\beta}^\star$ be optimal dual variables that solve the problem

$$d^\star \doteq \max_{\vec{\alpha},\vec{\beta}\geq\vec{0}} g(\vec{\alpha},\vec{\beta}). \tag{60}$$

It turns out that $\vec{\alpha}^\star$ can also be obtained by solving the quadratic program:

$$\begin{aligned} \min_{\vec{\alpha}} \quad & \left\|\vec{\alpha}+\vec{\beta}^\star\right\|_2^2 \\ \text{s.t.} \quad & \vec{\alpha}\geq\vec{0}. \end{aligned} \tag{61}$$

Solve this quadratic program (61) directly and find $\vec{\alpha}^\star$.

HINT: The duality or KKT approaches are not recommended. Consider $\vec{\alpha} = \begin{bmatrix}\alpha_1 & \cdots & \alpha_n\end{bmatrix}^\top$, and use the components of $\vec{\alpha}$ to decompose the problem into n separate scalar problems. Solve each one by checking critical points; that is, points where the gradient is 0, the boundary of the feasible set, and $\pm\infty$.

Solution: We have that

$$\left\|\vec{\alpha}+\vec{\beta}^\star\right\|_2^2 = \sum_{i=1}^{n}(\alpha_i+\beta_i^\star)^2. \tag{62}$$

Also, the $\vec{\alpha}\geq\vec{0}$ constraint is n separate constraints of the form $\alpha_i \geq 0$. Thus, we can solve for each α_i separately as

$$\alpha_i^\star \in \operatorname*{argmin}_{\alpha_i\geq 0}(\alpha_i+\beta_i^\star)^2. \tag{63}$$

This problem is convex and so we can solve it by checking the critical points.

- The gradient (w.r.t. α_i) is 0 if and only if $\alpha_i = -\beta_i^\star$. If $\beta_i^\star > 0$ then this solution is infeasible, and if $\beta_i^\star = 0$ then $\alpha_i = 0$.
- The constraint boundary is $\alpha_i = 0$; this solution is feasible with objective value $(\beta_i^\star)^2$.
- The limit $\alpha_i \to +\infty$ makes the objective value arbitrarily large, much larger than $(\beta_i^\star)^2$. The limit $\alpha_i \to -\infty$ makes the solution infeasible.

Thus the optimal solution for each scalar problem is $\alpha_i^\star = 0$. Thus $\vec{\alpha}^\star = \vec{0}$.

(i) Let $\beta^\star$ be a solution to the dual problem. Characterize the pairs $(\vec{x}_i, y_i)$ which are "support vectors", i.e., contribute to the optimal weight vector $\vec{w}^\star$, in terms of $\beta^\star$.

Solution: We have that $\vec{w}^\star = \sum_{i=1}^{n}\beta_i^\star\vec{z}_i$. If $\beta_i^\star > 0$ then $\vec{z}_i$ contributes to $\vec{w}^\star$, so $(\vec{x}_i, y_i)$ is a support vector. Otherwise $\beta_i^\star = 0$, then $\vec{z}_i$ does not contribute to $\vec{w}^\star$, so $(\vec{x}_i, y_i)$ is not a support vector.

7. Homework Process

With whom did you work on this homework? List the names and SIDs of your group members.

NOTE: If you didn't work with anyone, you can put "none" as your answer.

