CSE 203B W21 Midterm 10AM 2/16/2021-10AM 2/18/2021

Submit your solution to gradescope before the due time.

Policy of the Exam: Here is the policy of the exam:

1. This is an open-book take-home exam. Internet search is permitted. However, you are required to work by yourself. Consultation or discussion with any other parties is not allowed.
2. You are not required to typeset your solutions. We do expect your writing to be legible and your final answers clearly indicated. Also, please allow sufficient time to upload your solutions.
3. You are allowed to check your answers with programs in Matlab, CVX, Mathematica, Maple, NumPy, etc. Be aware that these programs may not produce the intermediate steps needed to receive credit.
4. If something is unclear, state the assumptions that seem most natural to you and proceed under those assumptions. Out of fairness, we will not be answering questions about the technical content of the exam on Piazza or by email. The solution will then be graded based on the reasonable assumptions made.

Part I: True or False: Explain your answer with one sentence (36 pts)

Problem 1 (convex set): Set $\{(x^3, y^2 - x)|x + 4y < 3, x, y \in R\}$ is convex.
T/F: **F**. We can check the convexity with a two dimensional plot.

Problem 2 (dual cone): Given cone
$K = \{\theta_1 u_1 + \theta_2 u_2 | u_1 = [2, -1]^T, u_2 = [1, 0]^T, \theta_1 \geq 0, \theta_2 \geq 0\}$, its dual cone is
$K^* = \{x_1 u_1 + x_2 u_2 | u_1 = [1, 2]^T, u_2 = [0, -1]^T, x_1 \geq 0, x_2 \geq 0\}$.
T/F: **T**. We can verify with a two dimensional plot.

Problem 3 (Convex Function): Given function $f(x, y) = x^T Ax + 2x^T By + y^T Cy$, where matrices $A, C \in S^n$ and $x, y \in R^n$, then $f(x, y)$ is concave if and only if the matrix $\begin{bmatrix} A & B \\ B^T & C \end{bmatrix}$ is negative semidefinite.
T/F: **T**. We can prove the statement via conversion to matrix formulation.

Problem 4 (Convex Function): Function $h(x) = \sum_{i=1}^m a_i f_i(x)$ is convex if function $f_i(x)$ is convex and $a_i \in R$ for every $i \in \{1, ..., m\}$.
T/F: **F**. The convexity holds only when a_is are positive.

Problem 5 (Convex Function): Function $h(x) = f(x) \times g(x)$ is convex if both functions $f(x)$ and $g(x)$ are convex.
T/F: **F**. We can show the result with counter examples.

Problem 6 (Quadratic Optimization Problem): A second-order cone programming (SOCP) problem is solved as a typical quadratically constrained quadratic programming (QCQP) problem.

T/F: **F**. The SOCP formulation covers a larger set of problems.

Problem 7 (Geometric Programming): A posynomial function is a convex function.
T/F: **F**. A posynomial function can be nonconvex.

Problem 8 (Duality): Given a nonconvex programming problem as a primal problem, the **dual of its dual** has the same solution as the primal problem.
T/F: **F**. The statement requires the convexity of the primal problem and the Slater condition.

Problem 9 (Duality): Given a function $f(x, y)$, the equality $\min_x \max_y f(x, y) = \max_y \min_x f(x, y)$ is always true.
T/F: **F**. The statement requires the existance of the saddle point.

Part II: Problem Solving: Show your process

Problem 1. Support vector machine (SVM): Given two sets of points $C = \{x_1, ..., x_m\}$ and $D = \{y_1, ..., y_m\}$, where $x_i, y_i \in R^n$, we find a hyperplane with vector $a \in R^n$ and bias $b \in R$ to maximize the following objective function. (20 pts)

$$\begin{aligned} &\max u, u \in R \\ \text{s.t. } &a^T x_i \leq b - u, \quad a^T y_i \geq b + u, \quad for\ all\ i = 1, ..., m \\ &||a||_2^2 \leq 1 \end{aligned}$$

(1) Given the above primal problem, formulate its dual problem.
(2) Create a nontrivial numerical example with $n = 2, m = 5$. Derive the solution via the primal formulation.
(3) Derive the solution of (2) via the dual formulation.
(4) Show the classification results of (2) and (3) with a 2-D plot (or plots).

Answer:

$$\begin{aligned} L(u, a, b, \lambda_1, \lambda_2, \lambda_3) &= -u + \lambda_1^T (Xa - (b - u)\mathbf{1}) + \lambda_2^T ((b + u)\mathbf{1} - Ya) + \lambda_3 (a^T a - 1) \\ &= -u(-1 + \lambda_1^T \mathbf{1} + \lambda_2^T \mathbf{1}) + b(-\lambda_1^T \mathbf{1} + \lambda_2^T \mathbf{1}) + (X^T \lambda_1 - Y^T \lambda_2)^T a + \lambda_3 a^T a - \lambda_3 \\ g(\lambda_1, \lambda_2, \lambda_3) &= \inf_{u,a,b} L(u, a, b, \lambda_1, \lambda_2, \lambda_3) \quad \text{unbounded if } \lambda_1^T \mathbf{1} + \lambda_2^T \mathbf{1} \neq 1 \text{ or } \lambda_1^T \mathbf{1} - \lambda_2^T \mathbf{1} \neq 0 \end{aligned}$$

$$\begin{aligned} \nabla_a L = 0 \implies & a = \frac{1}{2\lambda_3} (Y^T \lambda_2 - X^T \lambda_1) \\ \implies & g(\lambda_1, \lambda_2, \lambda_3) = -\frac{1}{4\lambda_3} ||X^T \lambda_1 - Y^T \lambda_2||_2^2 - \lambda_3 \text{ provided the above conditions are met.} \end{aligned}$$

So, the dual problem is

$$\max_{\lambda_1,\lambda_2} -||X^T\lambda_1 - Y^T\lambda_2||_2^2$$
$$\text{s.t. } \lambda_1^T\mathbf{1} = \lambda_2^T\mathbf{1} = \frac{1}{2}, \quad \lambda_1, \lambda_2 \geq 0$$

Note: I looked primarily for a correct Lagrangian and boundedness conditions. I also accepted reductions to the standard svm primal/dual problems.

Problem 2. Conjugate Function: Consider the function

$$f(x) = \begin{cases} ||x||_2^2, & ||x||_2 \leq a, \\ a(2||x||_2 - a), & ||x||_2 > a, \end{cases}$$

where variable $x \in R^n$ and constant $a \in R_{++}$. Derive the conjugate function $f^*(y)$, $y \in R^n$. (10 pts)

Answer:

$$f^*(y) = \begin{cases} ||y||_2^2/4, & \text{if } ||y||_2 \leq 2a, \\ \infty, & \text{otherwise.} \end{cases}$$

Problem 3. Kullback-Leibler Divergence: Show that $KL(p,q) = 0$ if and only if $p = q$. Here, we define the Kullback-Leibler (KL) divergence as $KL(p,q) = \sum_i p_i \log(\frac{p_i}{q_i})$, where $p_i > 0,\ q_i > 0,\ \forall i \in \{1,2,...,n\}$, and $\sum_{i=1}^n p_i = \sum_{i=1}^n q_i = 1$. (10 pts)

Answer:

Either way is correct:

1. Use Jensen's inequality and state the condition that the equality holds.
2. Show that the function is strictly convex and always larger than zero when $p \neq q$. Hence, proof by contradiction, $KL(p,q) = 0$ if and only if $p = q$.

Problem 4. Linear Programming Problem: Given the following optimization problem:

$$\begin{aligned} &\text{minimize } c^T x \\ \text{subject to} \quad & \\ &||x||_p = 1 \\ &-x \preceq 0 \end{aligned}$$

where variable $x \in R^n$, and constants $c \in R^n, p \geq 1$, derive an explicit solution. (10 pts)

Answer:

$$f^*(x) = \begin{cases} -||c||_q, & \text{if } c_i < 0\ \forall i; \\ -||c'||_q, & \text{if } \exists\, i,j \text{ s.t. } c_i < 0 \text{ and } c_j \geq 0, \\ & \text{Replace } c_j \text{ with } c'_j = 0 \text{ if } c_j \geq 0 \text{ else } c'_j = c_j\ \forall j \\ & \text{Set } c' \text{ be the vector of elements } c'_j. \\ c_{min}, & \text{if} c_i \geq 0\ \forall i,\ \text{where } c_{min} = \min_i c_i. \end{cases}$$

Problem 5. Compressed Sensing: Given the following problem

$$\text{minimize } ||y||_2^2 + \alpha||x||_1$$
$$\text{subject to}$$
$$\mathbf{A}x + y = b$$

where variables $x, y \in R^n$, matrix $A \in R^{m\times n}$, and constant $b \in R^n$. (14 pts)
(1) Write the dual formulation assuming that x_i is positive for all $i \in \{1, ..n\}$.
(2) Repeat item (1) without the assumption that x_i is positive.

Answer:

$$L(x, y, \upsilon) = ||y||_2^2 + \alpha||x||_1 + \langle \upsilon, Ax + y - b\rangle \qquad \text{Lagrangian}$$

$$g(x, y, \upsilon) = \inf_{x,y} \underline{||y||_2^2 + \langle \upsilon, y\rangle} + \underline{\alpha||x||_1 + \langle \upsilon, Ax\rangle} - \langle \upsilon, b\rangle$$

$$= \inf_{y} \underline{||y||_2^2 + \langle \upsilon, y\rangle} + \inf_{x} \underline{\alpha||x||_1 + \langle \upsilon, Ax\rangle} - \langle \upsilon, b\rangle$$

$$\inf_{y} \underline{||y||_2^2 + \langle \upsilon, y\rangle} \to y = -\frac{1}{2}\upsilon$$

$$\inf_{x} \underline{\alpha||x||_1 + \langle \upsilon, Ax\rangle} = 0 \text{ if } ||\upsilon^T A||_\infty \leq \alpha, -\infty \quad \text{o/w}$$

$$\text{since } \inf_{x_i} \underline{\alpha|x_i| + \langle \upsilon, Ax\rangle_i} = \inf_{x_i} \alpha|x_i| + [\upsilon^T Ax]_i \qquad \text{coordinate separability of } \ell_1 \text{ norm}$$

$$= \inf_{x_i} (\alpha + \text{sign}(x_i)(\upsilon^T A)_i)|x_i|$$

$$= 0 \quad \text{if } |\upsilon^T A|_i \leq \alpha, \quad -\infty \quad \text{o/w}$$

So, the dual problem for (2) ((2) generalizes (1)) is

$$\max_{\upsilon} -\frac{1}{4}||\upsilon||_2^2 - \upsilon^T b$$
$$\text{s.t. } ||\upsilon^T A||_\infty \leq \alpha$$

Note: no positivity constraints on the dual variables are necessary due to the primal equality constraint. I looked primarily for a correct lagrangian and upper bound constraints.