CME 307/MS&E 311: Optimization
Instructor: Prof. Yinyu Ye
Midterm Exam: Winter 2021-2022

HONOR CODE

In taking this examination, I acknowledge and accept Stanford University Honor Code.

NAME (Signed) : ______________________________

NAME (Printed): ______________________________

Problem	Full Points	Score
1	30	
2	20	
3	20	
4	20	
5	10	
6	20	
Bonus	10	
Total	120 + 10	

CME 307/MS&E 311 **Sample Midterm**
Optimization **Winter 2021-2022**
Prof.Yinyu Ye **24 hours**

Note: You have 24 hours to work on the exam once you click on the "Take the Quiz" link on Canvas. No late exams will be accepted. Upload your solutions back to the Quiz page. In taking this examination, you acknowledge and accept the Stanford University Honor Code.

Question 1 (30') [True/False] Give a *true* or *false* answer to each of the following questions and explain your choice. (If your answer is *true*, provide an argument or cite the appropriate claims from lecture notes or textbook. If your answer is *false*, provide a counterexample or cite the appropriate claims from lecture notes or textbook).

(a) (5') In linear programming, even if the problem is feasible and bounded, the optimal solution can be not achievable.

(b) (5') In conic linear programming, strong duality always holds if both of the primal and the dual problems are feasible. In contrast, in semi-definite programming, there are feasible problem pairs where strong duality does not hold.

(c) (5') In a general constrained optimization problem, if x is a local minimizer, then it is a KKT solution. However, the converse may not be true, that is, a KKT solution might not be a local minimizer of the problem.

(d) (5') Consider a set C defined by $C := \{\mathbf{x} \in \mathbb{R}^n_+ : f(\mathbf{x}) \geq 0\}$, where $n \geq 2$. If $f(\mathbf{x}) = \Pi_{i=1}^n x_i - 1$, then C is a convex set while $f(\mathbf{x})$ is a non-convex function on C.

(e) (5') Consider a conic LP in the standard form

$$\begin{aligned} \min_{\mathbf{x}} \quad & \mathbf{c} \bullet \mathbf{x} \\ \text{s.t.} \quad & \boldsymbol{a}_i \bullet \mathbf{x} = b_i, \ i = 1, ..., m, \ (\mathcal{A}\mathbf{x} = \mathbf{b}) \\ & \mathbf{x} \in K. \end{aligned}$$

Similar to the Lagrangian function, we can construct

$$L(\mathbf{x}, \mathbf{y}, \mathbf{s}) = \mathbf{c} \bullet \mathbf{x} - \mathbf{y}^T(\mathcal{A}\mathbf{x} - \mathbf{b}) - \mathbf{s} \bullet \mathbf{x},$$

where $\mathbf{y} \in \mathbb{R}^m$ and $\mathbf{s} \in K^*$. Let

$$\phi(\mathbf{y}, \mathbf{s}) = \inf_{\mathbf{x}: \mathcal{A}\mathbf{x} = \mathbf{b}, \mathbf{x} \in K} L(\mathbf{x}, \mathbf{y}, \mathbf{s}).$$

Then, the conic dual problem is equivalent to $\max_{\mathbf{y}, \mathbf{s}} \phi(\mathbf{y}, \mathbf{s})$.

(f) (5') Consider a feasible optimization problem on R^3,

$$\begin{aligned} \min \quad & f(x_1, x_2, x_3) \\ \text{s.t.} \quad & c_i(x_1, x_2, x_3) \leq 0, \quad i = 1, ..., m, \end{aligned}$$

where $f(x)$ and $c_i(x_1, x_2, x_3)$ are strictly convex functions[1], for all $i = 1, 2, 3, \ldots, m$. Assume that $f(x_1, x_2, x_3) = f(x_1, x_3, x_2)$ and $c_i(x_1, x_2, x_3) = c_i(x_1, x_3, x_2)$, for all i and any $x_1, x_2, x_3 \in R$. Assume this problem has one minimizer. Then we must have $x_2 = x_3$ at the unique minimizer.

[1] A strictly convex function f satisfies $f(\alpha x + (1-\alpha)y) < \alpha f(x) + (1-\alpha) f(y)$ for distinct x, y and $\alpha \in (0, 1)$

Question 2 [SVM] (20')

Recall the Supporting Vector Machine in lecture 1 and question 6 in Homework 2. Let the red class of points contain three points

$$\mathbf{a}_1 = (0, 3), \mathbf{a}_2 = (1, 2), \mathbf{a}_3 = (2, 3),$$

and let the blue class of points contain three points

$$\mathbf{b}_1 = (1, 0), \mathbf{b}_2 = (2, 1), \mathbf{b}_3 = (1, 3),$$

which are illustrated in Figure 1.

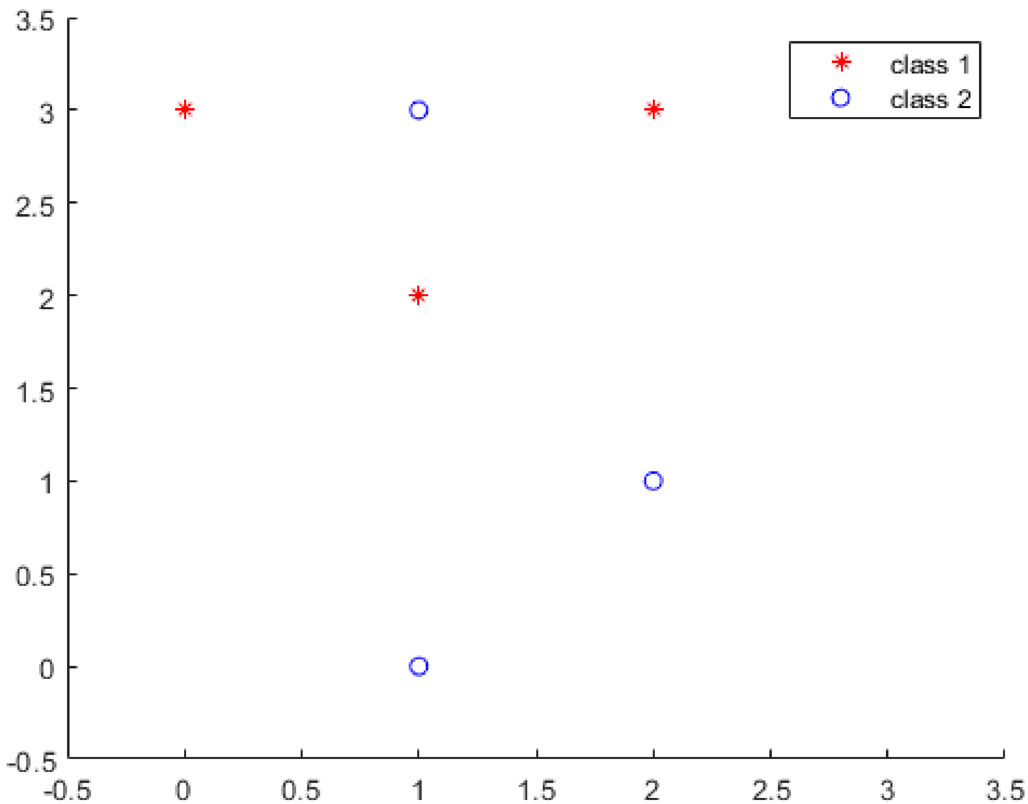

Figure 1: Scatter of all points

(a) (4') Consider the hard-margin SVM problem

$$\begin{aligned} \min \quad & \|\mathbf{x}\|^2 \\ \text{s.t.} \quad & \mathbf{a}_i^T\mathbf{x} + x_0 \geq 1, i = 1,2,3 \\ & \mathbf{b}_j^T\mathbf{x} + x_0 \leq -1, j = 1,2,3 \\ & \beta \geq 0. \end{aligned}$$

Is this problem feasible (i.e. there is a line that can separate the two classes of points)? Explain why.

(b) (6') Consider the following SVM problem with additional variable β (which was described in class):

$$\begin{aligned} \min \quad & \beta + \|\mathbf{x}\|^2 \\ \text{s.t.} \quad & \mathbf{a}_i^T\mathbf{x} + x_0 + \beta \geq 1, i = 1,2,3 \\ & \mathbf{b}_j^T\mathbf{x} + x_0 - \beta \leq -1, j = 1,2,3 \\ & \beta \geq 0. \end{aligned}$$

Show that this problem is feasible. Find an optimal solution by hand.

(c) (6') Consider the following soft-margin SVM problem with six additional variable β's:

$$\begin{aligned} \min \quad & \frac{1}{6}(\beta_1^a + \beta_2^a + \beta_3^a + \beta_1^b + \beta_2^b + \beta_3^b) + \|\mathbf{x}\|^2 \\ \text{s.t.} \quad & \mathbf{a}_i^T\boldsymbol{x} + x_0 + \beta_i^a \geq 1, i = 1,2,3 \\ & \mathbf{b}_j^T\boldsymbol{x} + x_0 - \beta_j^b \leq -1, j = 1,2,3 \\ & \beta_i^a, \beta_j^b \geq 0, \ i,j = 1,2,3. \end{aligned}$$

Construct the dual of the problem and find an optimal solution by hand or any optimization solver. (The optimal solution might not be unique and you only need to find one optimal solution for this question.)

If you choose to solve this question by hand, you can use the knowledge that the optimal solution has the six additional variables:

$$\beta_1^a = 0.51,\ \beta_2^a = 0.93,\ \beta_3^a = 0.68,$$
$$\beta_1^b = 0.42,\ \beta_2^b = 0.67,\ \beta_3^b = 1.42,$$

to simplify computation. Moreover, you can assume that the dual variables corresponding to the three constraints below are strictly positive:

$$\mathbf{a}_1^T \boldsymbol{x} + x_0 \geq 0.49$$
$$\mathbf{a}_3^T \boldsymbol{x} + x_0 \geq 0.32$$
$$\mathbf{b}_1^T \boldsymbol{x} + x_0 \leq -0.58$$

(If you choose to solve this question by any numerical solver, you don't need to use the above knowledge.)

(d) (4') Compare the results of part (a), (b) and (c). Explain which model would you likely to choose to use for this case and why. (There is no right or wrong answer to this question.)

Question 3 [Separable Nonlinear Programming] (20')

A function $f : R^n \longrightarrow R$ is said to be *separable* if it can be written in the form

$$f(x_1, \ldots, x_n) = \sum_{j=1}^{n} f_j(x_j).$$

Such functions were studied by Gibbs in connection with work on the chemical equilibrium problem (1876). He showed a theorem which states that (in the differentiable case) a necessary condition of local optimality for a feasible point x^* of the separable nonlinear programming problem

$$\begin{array}{ll} \text{minimize}_{\mathbf{x}} & \displaystyle\sum_{j=1}^{n} f_j(x_j) \\ \text{subject to} & \displaystyle\sum_{j=1}^{n} x_j = M \quad (M > 0) \\ & x_j \geq 0 \quad j = 1, \ldots, n \end{array}$$

is that there exist a thresholding number λ^* such that

$$f_j'(x_j^*) = \lambda^* \quad \text{if } x_j^* > 0$$

$$f_j'(x_j^*) \geq \lambda^* \quad \text{if } x_j^* = 0$$

Here f' is the derivative of f, and given M represents the total amount of the single resource. This constraint is called "Knapsack" constraint.

(a) (5') Show any feasible solution is a regular point so that every minimizer is an KKT solution (Hint: at lease one of variable must be positive at any feasible solution).

(b) (5') Why is this theorem true? What does λ^* represent?

(c) (5') Show that

$$\lambda^* = \frac{1}{M}\sum_j (x_j^* f_j'(x_j^*)).$$

(d) (5') What is λ^* if

$$f_j(x_j) = -w_j \log(x_j),\ \forall j,$$

where w_j is a given positive constant (the minimizer of the problem is called the weighted analytic center)?

Question 4 [Small SNL] (20')

In this problem, we will solve the SDP relaxation of a small SNL problem by hand. Consider a sensor network localization problem with three anchors:

$$
\begin{aligned}
\mathbf{a}_1 &= (1; 0), \\
\mathbf{a}_2 &= (0; 1), \\
\mathbf{a}_3 &= (-1; 0).
\end{aligned}
$$

The sensor's true location is

$$\mathbf{x} = (1; 1).$$

However, the sensor's true location is unknown to the problem solver. In Figure 2, the three red points denote anchors and the blue point denotes the true location of the sensor.

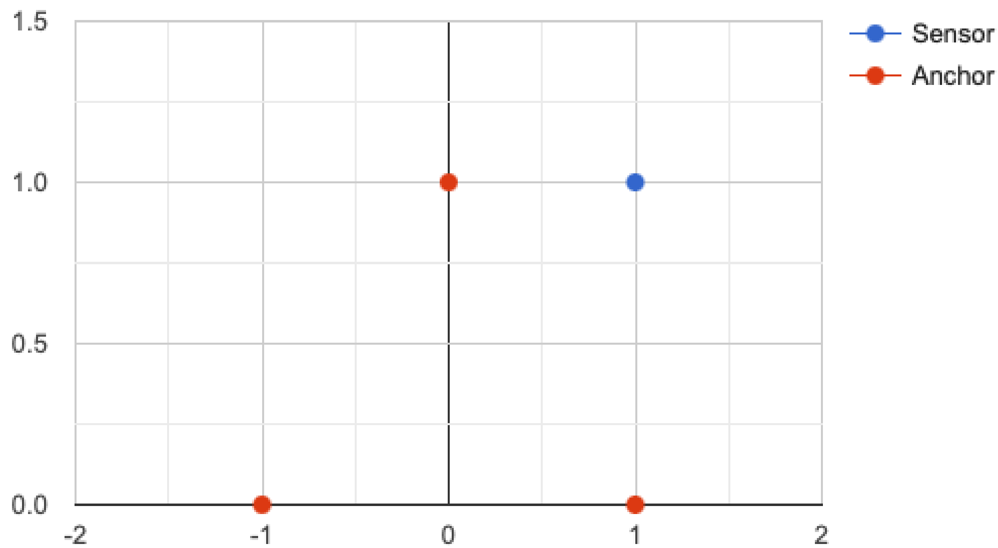

Figure 2: Location of anchors and the sensor

(a) (5') Write out the SDP relaxation problem for localizing the single sensor with null objective.

(b) (5') Write out the dual of the SDP relaxation problem.

(c) (5') Write out the SDP solution explicitly constructed from the true position (slide 12 of lecture note 5) and verify that it is an optimal solution to the SDP relaxation problem.

(d) (5') Show that the dual has a rank-one optimal solution so that the SDP solution in (c) is the only optimal solution.

Question 5 (10'+5'(bonus))

Consider the primal feasible region in standard form $\mathbf{Ax} = \mathbf{b},\ \mathbf{x} \geqslant \mathbf{0}$, where $\mathbf{A}$ is an $m \times n$ matrix, $\mathbf{b}$ is a constant nonzero m-vector, and $\mathbf{x}$ is a variable n-vector.

(a) (5') A variable x_i is said to be a *null variable* if $x_i = 0$ in every feasible solution. Prove that, if the feasible region is non-empty, x_i is a null variable if and only if there is a nonzero vector $\mathbf{y} \in \mathbb{R}^m$ such that $\mathbf{y}^T\mathbf{A} \geq \mathbf{0},\ \mathbf{y}^T\mathbf{b} = 0$. and the *ith* component of $\mathbf{y}^T\mathbf{A}$ is strictly positive.

(b) (5') [*Strict complementarity*] Let the feasible region be nonempty. Then there is a feasible $\mathbf{x}$ and vector $\mathbf{y} \in \mathbb{R}^m$ such that

$$\mathbf{y}^T\mathbf{A} \geq \mathbf{0},\ \mathbf{y}^T\mathbf{b} = 0,\ \mathbf{y}^T\mathbf{b} + \mathbf{x} > \mathbf{0}.$$

(c) (5') (Bonus) A variable x_i is a *nonextremal variable* if $x_i > 0$ in every feasible solution. Prove that, if the feasible region is non-empty, x_i is a nonextremal variable if and only if there is $\mathbf{y} \in \mathbb{R}^m$ and $\mathbf{d} \in \mathbb{R}^n$ such that $\mathbf{y}^T\mathbf{A} = \mathbf{d}^T$, where $d_i = -1,\ d_j \geqslant 0$ for $j \neq i$; and such that $\mathbf{y}^T\mathbf{b} < 0$.

Question 6 (20')

We have studied the Fisher market equilibrium problem where there are m goods in the market and each good i has a fixed amount $\bar{s}_i(> 0)$ available. There are n buyers in the market where each buyer, say buyer $j \in \{1, ..., n\}$, is equipped with a fixed budget $w_j(> 0)$ and independently solves a linear utility maximization problem

$$\max_{\mathbf{x}_j} \quad u_j(\mathbf{x}_j) = \sum_i u_{ij}x_{ij} \quad \text{s.t. } \mathbf{p}^T\mathbf{x}_j \leq w_j,\ \mathbf{x}_j \geq \mathbf{0}.$$

Here $u_{ij} \geq 0$ is the coefficient of buyer j on good i, and decision variable x_{ij} is the amount of good i purchased by buyer j, and $\mathbf{p} \in R^m$ is a given market price vector. The equilibrium prices are the prices to clear the market.

Another important utility function for each buyer j is called the Leontief utility function

$$u_j(\mathbf{x}_j) = \min_i \left\{\frac{x_{ij}}{u_{ij}}\right\},$$

that is, the utility function value is the smallest x_{ij}/u_{ij}, $i = 1, ..., m$. For simplicity, assume $u_{ij} > 0$ for all i and j.

One fact of the Leontief utility function is that a good may not be able to be cleared (or all purchased) in the market so that its price should be zero from the economic theory. Thus, we look the equilibrium price $\mathbf{p} \in R^m$ and allocation $\mathbf{x}_j \in R^m$, $j = 1, ..., n$, such that the following conditions are met:

$$\sum_j x_{ij} \leq \bar{s}_i, \quad \text{and} \quad p_i\left(\bar{s}_i - \sum_{j=1}^n x_{ij}\right) = 0,\ \forall i,$$

and $\mathbf{x}_j \in R^m$, $j = 1, ..., n$, is an optimal solution of

$$\max_{\mathbf{x}_j} \quad \min_i \left\{\frac{x_{ij}}{u_{ij}}\right\} \quad \text{s.t. } \mathbf{p}^T\mathbf{x}_j \leq w_j,\ \mathbf{x}_j \geq \mathbf{0}. \tag{1}$$

(a) (5') What is the interpretation of the Leontief utility function?

(b) (5') Write down the optimality conditions of buyer jth's maximization problem (1). (Hint: you may simplify the problem by define a scalar utility objective value z_j, then $x_{ij} = u_{ij} z_j$ so that the problem become a one-variable problem.)

(c) (5') Derive the equilibrium (price $\mathbf{p}$ and allocation $\mathbf{x}_j,\ j = 1, ..., n$) conditions for the Leontoef market.

(d) (5') Derive a social optimization problem to represent these conditions.

Question 7 (Bonus 5')

Consider the unconstraint minimization problem:

$$\min_{\mathbf{x}\in\mathbb{R}^n} \quad f(\mathbf{x}) = \frac{1}{2}\mathbf{x}^\top \mathbf{Q}\mathbf{x} - \boldsymbol{b}^T\mathbf{x},$$

where $\mathbf{Q} \in \mathbb{R}^{n\times n}$ is a PSD matrix and $\boldsymbol{b} \in \mathbb{R}^n$ is a fixed vector. Assume $\mathbf{Q}$ has K distinct positive eigenvalues and denote them by $\lambda_1, ..., \lambda_K$.

Let $\mathbf{x}_0$ be an arbitrary vector in $\mathbb{R}^n$, and

$$\mathbf{x}_k = \mathbf{x}_{k-1} - \frac{1}{\lambda_k}\nabla f(\mathbf{x}_{k-1}),$$

for $k = 1, ..., K$. Show that if the optimization problem is bounded, $\mathbf{x}_K$ is one optimal solution.