# Convex Optimization – Midterm Exam

## Spring 2023

## 48 hour take-home

## Available during Tuesday 03/14, 2 pm to Friday 03/17, 5 pm

This exam is a 48-hour take-home exam. You can start anytime during the time window above. Once started, you will have 48 hours to complete, or until the time window expires, whichever is earlier. That is, if you start less than 48 hours from the ending time, you will only have until the ending time of the exam period to complete. So please start early.

This exam is open-book open-note, you can consult the textbook, lecture notes, and homework solutions, but no other materials should be used and browsing the internet for the exam solutions is not allowed. We take academic integrity seriously, please note that even just consulting the solutions to exam questions, if available elsewhere, is a violation. *Your turned-in exams should be your own work, keep in mind we give partial credits.* If you reference a result from the textbook, clearly write the page for the reference.

Please write your answers neatly in your paper and make sure that the scan or pictures are clearly legible (no blurry text please, it hurt our eyes and we can't understand what you write!).

All questions have an equal number of points but are not at the same level of difficulty. You need to answer 5 required questions, plus 1 choice question. Good luck!

*Note: Each question carries an equal number of 20 points. These 20 points will be divided among the sub-questions.*

1. *Identifying convex sets*

   (a) Is the following set convex? What type of set is it? Provide reasoning.

   $$S = \{x \in \mathbf{R}^n \mid x^T y \leq 1, |y_i| \leq 1 \text{ for } i = 1, ...n\}$$

   (b) Is the following set convex: The set of points whose distance to point $a$ does not exceed a fixed fraction $\theta$ of the distance to point $b$, i.e., the set

   $$\{x \mid ||x - a||_2 \leq \theta ||x - b||_2\}$$

   You can assume $a \neq b$ and $0 \leq \theta \leq 1$. Identify the set when $\theta = 1$ and when $0 < \theta < 1$.

2. *Identifying convex functions.* Determine if the following functions are convex, concave, quasiconvex, or quasiconcave. Show your arguments in detail (brief arguments without details will not receive full mark).

   (a) $f(x, y) = ||Ax - b||^2/y$, for $x \in \mathbf{R}^n$, $y \in \mathbf{R}_{++}$.

   (b)
   $$f(x) = \frac{||Ax - b||_2^2}{1 - x^T x}$$

   on $\{x \in \mathbf{R}^n \mid ||x||_2 < 1\}$. (Hint: Examine the epigraph and use the result in part (b).)

3. (a) *Augmenting features with the average.* You are given a set of input data $x_1, x_2, ..., x_m$, where each vector $x_i \in R^n$, and a set of measurement output $\hat{y} \in R^m$. Using these data, you are fitting a regression model $\hat{y} = x^T \beta + v$ to data, computing the model coefficients $\beta$ and $v$ using least squares. Formulate the least-square problem in a compact form. Is this problem convex? (Hint: You may want to form a matrix of the input data.)

   A friend suggests adding a new feature, which is the average of the original features. That is, he suggests using the new feature vector $\tilde{x} = (x, \mathbf{avg}(x))$. He explains that by adding this new feature, you might end up with a better model, so that you can estimate the coefficients $\beta$ and $v$ more accurately. (Of course, you would test the new model using validation.) Do you think this is a good idea? Provide reasoning if possible.

   Hint: Reformulate the least square problem with the new feature as an additional input.

(b) *Multiple channel marketing campaign*. Potential customers are divided into $m$ market segments, which are groups of customers with similar demographics, $e.g.$, college educated women aged 25–29. A company markets its products by purchasing advertising in a set of $n$ channels, $i.e.$, specific TV or radio shows, magazines, web sites, blogs, direct mail, and so on. The ability of each channel to deliver impressions or views by potential customers is characterized by the $reachmatrix$, the $m \times n$ matrix $R$, where $R_{ij}$ is the number of views of customers in segment $i$ for each dollar spent on channel $j$. (We assume that the total number of views in each market segment is the sum of the views from each channel, and that the views from each channel scale linearly with spending.) The $n$-vector $c$ will denote the company's purchases of advertising, in dollars, in the n channels. The $m$-vector $v$ gives the total number of impressions in the $m$ market segments due to the advertising in all channels. Finally, we introduce the $m$-vector $a$, where $a_i$ gives the profit in dollars per impression in market segment $i$. The entries of $R$, $c$, $v$, and $a$ are all nonnegative.

Formulate a problem to maximize the total profit subject to a total cost constraint. What type of problem is this, and is the problem convex?

4. *Fun with linear programming.* Consider the following linear programming problem:

$$\begin{aligned} &\text{minimize } c^T x + d \\ &\text{subject to } Ax \leq b \end{aligned}$$

where $x \in \mathbf{R}^2$ and

$$c = \begin{bmatrix} 15 \\ 1 \end{bmatrix}, \ d = 3, \ A = \begin{bmatrix} -6 & -20 \\ 22 & 1 \\ -24 & 2 \\ -3 & 1 \\ 8 & -5 \end{bmatrix}, \ b = \begin{bmatrix} 60 \\ 35 \\ 100 \\ 40 \\ 15 \end{bmatrix}$$

(a) Identify the feasible set $\mathcal{F}$ and sketch it. You can use a program to fo the plot if you wish. What type of set is it? (Hint: Suggested range for the horizontal axis of $x_1$ is $[-10, 10]$)

(b) On the same plot as (a), sketch the level sets of the objective function

$$L_\alpha = \{x \mid c^T x + d = \alpha\}$$

for $\alpha$ from $-100$ to $100$ in steps of 50. Based on that, find the optimal solution $x^*$ to the linear programming problem. Sketch the level set $L_{p^*}$ that achieves the optimal solution. What is the optimal value $p^*$?

(c) Now suppose that we add an additional constraint to the original LP as

$$g^T x = h$$

where

$$g = \begin{bmatrix} 20 \\ 2 \end{bmatrix}, \; h = -36,$$

Find the optimal solutions $(x^*, p^*)$ to this new problem. Show your work.

5. *Planning production with uncertain demand.* You must order (nonnegative) amounts $r_1, ..., r_m$ of raw materials, which are needed to manufacture (nonnegative) quantities $q_1, ..., q_n$ of $n$ different products. To manufacture one unit of product $j$ requires at least $A_{ij}$ units of raw material $i$, so we must have $r \succeq Aq$. (We will assume that $A_{ij}$ are nonnegative.) The per-unit cost of the raw materials is given by $c \in \mathbf{R}_+^m$, so the total raw material cost is $c^T r$.

   The (nonnegative) demand for product $j$ is denoted $d_j$; the number of units of product $j$ sold is $s_j = \min\{q_j, d_j\}$. (When $q_j > d_j$ , $q_j - d_j$ is the amount of product $j$ produced, but not sold; when $d_j > q_j$, $d_j - q_j$ is the amount of unmet demand.) The revenue from selling the products is $p^T s$, where $p \in \mathbf{R}_+^n$ is the vector of product prices. The profit is $p^T s - c^T r$. (Both $d$ and $q$ are real vectors; their entries need not be integers.)

   You are given $A$, $c$, and $p$. The product demand, however, is not known. Instead, a set of $K$ possible demand vectors, $d^{(1)}, ..., d^{(K)}$, with associated probabilities $\pi_1, ..., \pi_K$, is given. (These satisfy $\mathbf{1}^T \pi = 1$, $\pi \succeq 0$.)

   You will explore two different optimization problems that arise in choosing $r$ and $q$ (the variables) as follows.

   (a) **Choose $r$ and $q$ ahead of time.** You must choose $r$ and $q$, knowing only the data listed above. (In other words, you must order the raw materials, and commit to producing the chosen quantities of products, before you know the product demand.) The objective is to maximize the expected profit.

   (b) **Choose $r$ ahead of time, and $q$ after $d$ is known.** You must choose $r$, knowing only the data listed above. Some time after you have chosen $r$, the demand will become known to you. This means that you will find out which of the $K$ demand vectors is the true demand. Once you know this, you must choose the quantities to be manufactured. (In other words, you must order the raw materials before the product demand is known; but you can choose the mix of products to manufacture after you have learned the true product demand.) The objective is to maximize the expected profit. (Hint: Think how you may introduce new production variables in this case which depends on the demand $d$.)

   Explain how to formulate each of these problems as a convex optimization problem. **Clearly state what the variables are in the problem, what the constraints are, and describe the roles of any auxiliary variables or constraints you introduce.** Which formulation do you think will lead to a higher profit and why? Is this obvious or can be inferred from the formulations themselves (without the word descriptions)?

# Choose one (and only one) question from the next 3 questions.

6. For a symmetric $n \times n$ matrix $A$, we define $f(A)$ as the optimal value of the semidefinite program

$$\begin{array}{ll} \text{minimize} & \mathbf{tr} X + \mathbf{tr} Y \\ \text{subject to} & \begin{bmatrix} X & A \\ A & Y + I \end{bmatrix} \succeq 0 \\ & Y \succeq 0 \end{array}$$

with variables $X \in \mathbf{S}^n$ and $Y \in \mathbf{S}^n$. Is $f(A)$ a convex function of $A$? Provide reasoning. (Hint: Try using a simple approach (for example, using the definition or convexity preservation operations), do not over-complicate the problem.

7. Show that $X = B^T A^{-1} B$ solves the SDP

$$\begin{array}{ll} \text{minimize} & \mathbf{tr} X \\ \text{subject to} & \begin{bmatrix} A & B \\ B^T & X \end{bmatrix} \succeq 0, \end{array}$$

with variable $X \in \mathbf{S}^n$, where $A \in \mathbf{S}^m_{++}$ and $B \in \mathbf{R}^{m \times n}$ are given. Conclude that $\mathbf{tr}(B^T A^{-1} B)$ is a convex function of $(A, B)$, for $A$ positive definite. Provide reasoning. (Hint: Use the Schur complement.)

8. Show that the following functions $f : \mathbf{R}^n \to \mathbf{R}$ are convex.

   (a) $f(x) = -\exp(-g(x))$ where $g : \mathbf{R}^n \to \mathbf{R}$ has a convex domain and satisfies

   $$\begin{bmatrix} \nabla^2 g(x) & \nabla g(x) \\ \nabla g(x)^T & 1 \end{bmatrix} \succeq 0$$

   for $x$ $\mathbf{dom}$ $g$.

   (b) The function

   $$f(x) = \max\{||APx - b|| \mid P \text{ is a permutation matrix}\}$$

   with $A \in \mathbf{R}^{m \times n}$, $b \in \mathbf{R}^m$. (Hint: Using simple rules which preserve function convexity.)