Due to the current COVID-19 pandemic, the final exam will be a take home final. In addition, your final course grade will be computed in a manner that deviates from the description given in the course syllabus. The final exam is now worth 20% of your grade and the midterm exam is worth 45% of your grade:

$$\text{quiz } 350 \; + \; \text{midterm } 450 \; + \; \text{final } 200 \; = \; 1000 \text{ points.}$$

No person-to-person contact is permitted for either retrieving or submitting the final exam.

I will email all students a PDF of the final exam at 5pm on Sunday evening March 15. A PDF, scan, or photocopy of your worked exam is to be emailed back to me by 10am on Monday March 16.

The exam will be composed of 3 sections:
I Quadratic Optimization and Least Squares
II Optimality Conditions
III Convex Optimization
The number of questions and points for each section varies with an emphasis on Section III Convex Optimization since this is the material we studied most intensively since the midterm exam. A more detailed description of the material you will be tested on in each area is given below. This will be followed by sample questions for each area. The number of sample questions greatly exceeds the number of questions on the exam.

I Quadratic Optimization and Least Squares

The questions for this section are of three types: theory, linear algebra and computation. Theory questions refer to the basic theorems and properties of quadratic optimization and linear least squares problems. In particular, you should know how quadratic optimization and linear least squares problems are related, when they differ, when they are the same, and when solutions exist, what is the structure of the solutions sets, and when are solutions unique. The linear algebra questions refer to the linear algebraic properties of the objective functions and the matrix factorizations associated with computing solutions. The computational questions require you to compute the solution set to concrete problems of this type which includes quadratic optimization problems with affine constraints.

II Optimality Conditions

These question relate to first- and second-order optimality conditions for both constrained and unconstrained problems with particular emphasis on NLP problems in the constrained case. You will need to know what a tangent cone is and be able to compute one (under the assumption of regularity). You will **not** need to work with constraints qualifications other than to know that they imply regularity. You will be asked to to answer questions on the theory (statements of theorems), testing optimality, computing solutions, and computing KKT points.

III Convex Optimization

For these questions you will need to know the definitions of convex functions and sets and be able to show that a given function or set is convex. You will need to know the optimality conditions for convex optimization problems especially for convex NLP. You will need to be able to compute KKT points. You will also need to know the rudiments of Lagrangian duality and how to compute the Lagrangian dual for simple problems.

Sample Questions

I Quadratic Optimization and Least Squares

(Questions taken from the midterm guide.)

1. Let $A \in \mathbb{R}^{m\times n}$ and $b \in \mathbb{R}^m$, and consider the linear least squares problem

$$\mathcal{LLS} \qquad \min \frac{1}{2}\|Ax - b\|_2^2 \ .$$

   a. Show that the matrix A^TA is always positive semi-definite and provide necessary and sufficient condition on A under which A^TA is positive definite.

   b. Show that $\mathrm{Nul}(A^TA) = \mathrm{Nul}(A)$ and $\mathrm{Ran}(A^TA) = \mathrm{Ran}(A^T)$.

   c. Show that $\mathcal{LLS}$ always has a solution.

   d. State and prove a necessary and sufficient condition on the matrix $A \in \mathbb{R}^{m\times n}$ under which $\mathcal{LLS}$ has a unique global optimal solution.

   e. Decribe the QR-factorization of A and show how it can be used to construct a solution to $\mathcal{LLS}$.

   f. How can the QR factorization of A be used to obtain the orthoronal projection onto $\mathrm{Ran}(A)$.

   g. If $\mathrm{Nul}(A) = \{0\}$, show that $(A^TA)^{-1}$ is well defined and that $P = A(A^TA)^{-1}A^T$ is the orthogonal projection onto $\mathrm{Ran}(A)$ with

$$\frac{1}{2}\|(I-P)b\|_2^2 = \min \frac{1}{2}\|Ax - b\|_2^2 \ .$$

   h. Let $A \in \mathbb{R}^{m\times n}$ be such that $\mathrm{Ran}(A) = \mathbb{R}^m$ and set $P = A^T(AA^T)^{-1}A$. Show that $Px^0 = A^T(AA^T)^{-1}b$ for every $x^0 \in \mathbb{R}^n$ satisfying $Ax^0 = b$ and that $\hat{x} := A^T(AA^T)^{-1}b$ is the unique solution to the problem

$$\min \frac{1}{2}\|x\|_2^2 \quad \text{subject to} \quad Ax = b \ .$$

2. Consider the matrix and vector

$$A = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 0 \\ 1 & 0 & 0 \\ 1 & 0 & 1 \end{bmatrix} \quad \text{and} \quad b = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix} .$$

   a. Compute the orthogonal projection onto $\mathrm{Ran}(A)$.

   b. Compute the orthogonal projection onto $\mathrm{Null}(A^T)$.

   c. Compute the QR factorization of A.

   d. Compute the LU factorization of A^TA.

   e. Compute the solution for the LLS problem for this matrix and vector using the LU factorization of A^TA.

   f. Compute the solution for the LLS problem for this matrix and vector using the QR factorization of A.

3. Consider the matrix and vector

$$A = \begin{bmatrix} 1 & -1 & 0 \\ 1 & 1 & 2 \\ 1 & -1 & 0 \\ 1 & 1 & 2 \end{bmatrix} \quad \text{and} \quad b = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 1 \end{pmatrix} .$$

a. Compute the orthogonal projection onto $\mathrm{Ran}(A)$.

b. Compute the orthogonal projection onto $\mathrm{Null}(A^T)$.

c. Compute the QR factorization of A.

d. Compute the LU factorization of $A^T A$.

e. Compute the solution for the LLS problem for this matrix and vector using the LU factorization of $A^T A$.

f. Compute the solution for the LLS problem for this matrix and vector using the QR factorization of A.

4. Let $a \in \mathbb{R}$ and consider the function

$$f(x_1, x_2, x_3) = \frac{1}{2}[x_1^2 + (x_1 - 2a^4)^2 + (x_1 - x_2)^2 + (ax_2 + x_3)^2].$$

(a) Write this function in the form of the objective function for a linear least squares problem by specifying the matrix A and the vector b.

(b) Describe the solution set of this linear least squares problem as a function of a.

5. Find the quadratic polynomial $p(t) = x_0 + x_1 t + x_2 t^2$ that best fits the following data in the least-squares sense:

$$\begin{array}{c|ccccc} t & -2 & -1 & 0 & 1 & 2 \\ \hline y & 2 & -10 & 0 & 2 & 1 \end{array}.$$

6. Consider the LLS problem with

$$A = \begin{bmatrix} 1 & -1 & 0 \\ 1 & 1 & 2 \\ 1 & -1 & 0 \\ 1 & 1 & 2 \end{bmatrix} \quad \text{and} \quad b = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 0 \end{bmatrix}.$$

(a) What are the normal equations for this A and b.

(b) Solve the normal equations to obtain a solution to the problem LLS for this A and b.

(c) What is the general reduced QR factorization for the matrix A?

(d) Compute the orthogonal projection onto the range of A.

(e) Use the recipe given in the notes for solving the LLS problem associated with this A and b using the QR-factorization you have computed for A.

(f) If $\bar{x}$ solves LLS for this A and b, what is $A\bar{x} - b$?

7. Consider the function

$$f(x) = \frac{1}{2}x^T Q x + c^T x,$$

where $Q \in \mathbb{R}^{n \times n}$ is symmetric and $c \in \mathbb{R}^m$.

(a) Give necessary and sufficient conditions on Q and c for which there exists a solution to the problem $\min_{x \in \mathbb{R}^n} f(x)$.

(b) If Q is positive definite, show that there is a nonsingular matrix L such that $Q = LL^T$.

(c) With Q and L as defined in the part (2), show that

$$f(x) = \frac{1}{2}\|L^T x + L^{-1}c\|_2^2 - \frac{1}{2}c^T Q^{-1} c.$$

(d) If Q is psd, under what conditions is $\min_{x \in \mathbb{R}^n} f(x) = -\infty$?

(e) Let $\hat{x} \in \mathbb{R}^n$ and S be a subspace of $\mathbb{R}^n$. Give necessary and sufficient conditions on Q and c for which there exists a solution to the problem

$$\min_{x \in \hat{x}+S} f(x) \ .$$

(f) Show that every local solution to the problem $\min_{x \in \mathbb{R}^n} f(x)$ is necessarily a global solution.

(g) Let $A \in \mathbb{R}^{m \times n}$ and $b \in \mathbb{R}^m$ and consider the problem $\min \left\{ \frac{1}{2} x^T Q x + c^T x \,|Ax = b \right\}$, where it is assumed that the system $Ax = b$ is consistent.

i. Using a Lagrange multiplier vector $y \in \mathbb{R}^m$ give a necessary and sufficient condition under which $\bar{x} \in \mathbb{R}^m$ is and optimal solution to this problem.

ii. Under what conditions is $(\bar{x}, \bar{y})$ the unique solution and Lagrange multipier pair for this problem.

iii. Provide a necessary and sufficient condition under which $\bar{x}$ is a unique solution to this problem.

(h) Compute the Cholesky factorizations of the following matrices.

$$(a)\ H = \begin{bmatrix} 2 & 1 & 0 \\ 1 & 2 & 1 \\ 0 & 1 & 2 \end{bmatrix} \qquad (b)\ H = \begin{bmatrix} 2 & 1 & 0 \\ 1 & 2 & 1 \\ 0 & 1 & -2 \end{bmatrix}$$

$$(c)\ H = \begin{bmatrix} 5 & 2 & -1 \\ 2 & 1 & -1 \\ -1 & -1 & 2 \end{bmatrix} \qquad (d)\ H = \begin{bmatrix} 1 & 4 & 1 \\ 4 & 20 & 2 \\ 1 & 2 & 2 \end{bmatrix} .$$

8. For each of the matrices H and vectors g below, determine the optimal value in

$$\mathcal{Q} : \min_x f(x) := \frac{1}{2} x^T H x + g^T x.$$

If an optimal solution exists, compute the complete set of optimal solutions.

a.
$$H = \begin{bmatrix} 2 & 1 & 0 \\ 1 & 2 & 1 \\ 0 & 1 & 2 \end{bmatrix} \quad \text{and} \quad g = \begin{bmatrix} 3 \\ 1 \\ 1 \end{bmatrix} .$$

b.
$$H = \begin{bmatrix} 2 & 1 & 0 \\ 1 & 2 & 1 \\ 0 & 1 & -2 \end{bmatrix} \quad \text{and} \quad g = \begin{bmatrix} 3 \\ 1 \\ 1 \end{bmatrix} .$$

c.
$$H = \begin{bmatrix} 5 & 2 & -1 \\ 2 & 1 & -1 \\ -1 & -1 & 2 \end{bmatrix} \quad \text{and} \quad g = \begin{bmatrix} 3 \\ 1 \\ 0 \end{bmatrix} .$$

9. Consider the matrix $H \in \mathbb{R}^{3 \times 3}$ and vector $g \in \mathbb{R}^3$ given by

$$H = \begin{bmatrix} 1 & 4 & 1 \\ 4 & 20 & 2 \\ 1 & 2 & 2 \end{bmatrix} \quad \text{and} \quad g = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} .$$

Does there exists a vector $u \in \mathbb{R}^3$ such that $f(tu) \xrightarrow{t \uparrow \infty} -\infty$? If yes, construct u.

10. Consider the linearly constrained quadratic optimization problem

$$\mathcal{Q}(H,g,A,b) \qquad \begin{array}{ll} \text{minimize} & \frac{1}{2}x^T Hx + g^T x \\ \text{subject to} & Ax = b\ , \end{array}$$

where H, A, g, and b are given by

$$H = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 2 & 1 \\ 0 & 1 & 3 \end{bmatrix},\ g = (1,1,1)^T,\ b = (4,2)^T, \quad \text{and} \quad A = \begin{bmatrix} 1 & 2 & 1 \\ 1 & 0 & 1 \end{bmatrix}.$$

a. Compute a basis for the null space of A.

b. Solve the problem $\mathcal{Q}(H,g,A,b)$ two different ways. One should use a basis for the null space of A and the other should not.

11. Let $H \in \mathbb{R}^{n\times n}$ be symmetric and positive definite, $r \in \mathbb{R}^n \setminus \mathrm{Span}[\mathsf{e}]$, $\mu \in \mathbb{R}$. Solve the problem

$$\begin{array}{ll} \text{minimize} & \frac{1}{2}x^T Hx \\ \text{subject to} & \mathsf{e}^T x = 1 \quad \text{and} \quad r^T x = \mu\ , \end{array}$$

where $\mathsf{e} := (1,1,\dots,1)^T \in \mathbb{R}^n$ is the vector of all ones and $\mu \in \mathbb{R}$.

II Optimality Conditions

1. State the first- and second-order conditions for optimality for the following three problems:

   (a) Linear least squares: $\min_{x\in\mathbb{R}^n} \frac{1}{2}\|Ax-b\|_2^2$, where $A \in \mathbb{R}^{m\times n}$ and $b \in \mathbb{R}^m$.

   (b) Quadratic Optimization: $\min_{x\in\mathbb{R}^n} \frac{1}{2}x^T Qx + g^T x$, where $Q \in \mathbb{R}^{n\times n}$ is symmetric and $g \in \mathbb{R}^n$.

   (c) Quadratic Optimization with Affine Constraints: $\min_{x\in\mathbb{R}^n} \left\{\frac{1}{2}x^T Qx + g^T x \mid Ax = n\right\}$, where $Q \in \mathbb{R}^{n\times n}$ is symmetric, $A \in \mathbb{R}^{m\times n}$, $g \in \mathbb{R}^n$, and $b \in \mathbb{R}^m$.

2. Consider the nonlinear progamming (NLP) problem

$$\min_{x\in\Omega} f(x), \qquad (\spadesuit)$$

where

$$\Omega := \{x : c_i(x) \le 0, i = 1,\dots,s, c_i(x) = 0, i = s+1,\dots,m\}. \qquad (\clubsuit)$$

   (a) State the first-order optimally conditions for NLP.

   (b) What does it mean for $(\bar{x},\bar{y}) \in \mathbb{R}, \times\mathbb{R}^m$ to be KKT pair for NLP?

3. A critical point of a function $f : \mathbb{R}^n \to \mathbb{R}$ is any point x at which $\nabla f(x) = 0$. Compute all of the critical points of the following functions. Classify the critical points as local/global minima/maxima/saddle points. If no critical point exists, explain why.

   (a) $f(x) = x_1^2 - 4x_1 + 2x_2^2 + 7$

   (b) $f(x) = e^{-\|x\|^2}$

   (c) $f(x) = x_1^2 - 2x_1x_2 + \frac{1}{3}x_2^3 - 8x_2$

   (d) $f(x) = (2x_1 - x_2)^2 + (x_2 - x_3)^2 + (x_3 - 1)^2$

   (e) $f(x_1,\, x_2) := x_1^2 + 2x_2x_2 + x_2^2 - 8x_1 - 8x_2$

4. Compute the KKT points for the following problems.

   (a) $\min\left\{\frac{1}{2}\|x\|_2^2 \;\middle|\; a^T x \le \alpha\right\}$ where $a \in \mathbb{R}^n \setminus \{0\}$ and $\alpha \in \mathbb{R}$.
   (b) $\min\left\{\frac{1}{2}\|x\|_2^2 \;\middle|\; a^T x = \alpha,\ b^T x \le \beta\right\}$ where $a, b \in \mathbb{R}^n \setminus \{0\}$ with a and b linearly independent and $\alpha, \beta \in \mathbb{R}$.
   (c) $\min\left\{\frac{1}{2}\|x\|_2^2 \;\middle|\; a^T x \le \alpha,\ b^T x \le \beta\right\}$ where $a, b \in \mathbb{R}^n \setminus \{0\}$ with a and b linearly independent and $\alpha, \beta \in \mathbb{R}$.
   (d) $\min\left\{\frac{1}{2}x^T H x \;\middle|\; \mathsf{e}^T x = 1 \quad \text{and} \quad r^T x = \mu\right\}$ where $H \in \mathbb{S}^n_{++}$, $\mathsf{e} := (1, 1, \ldots, 1)^T \in \mathbb{R}^n$ is the vector of all ones, e and $r \in \mathbb{R}^n$ are linearly independent, and $\mu \in \mathbb{R}$.
   (e) $\min\left\{\frac{1}{2}x^T H x \;\middle|\; \mathsf{e}^T x = 1 \quad \text{and} \quad r^T x \ge \mu\right\}$ where $H \in \mathcal{S}^n_{++}$, $\mathsf{e} := (1, 1, \ldots, 1)^T \in \mathbb{R}^n$ is the vector of all ones, e and $r \in \mathbb{R}^n$ are linearly independent, and $\mu \in \mathbb{R}$.
   (f) $\min\left\{4x_1^2 + 4x_1x_2 + x_2^2 - 8x_1 - 4x_2 \;\middle|\; x_1 + x_2 \le 0,\ x_1 - x_2 \le 0\right\}$

III Convexity

1. Which of the following functions are convex? Which are strictly convex?

   (a) $f(x) = x_1^2 - 4x_1 + 2x_2^2 + 7$
   (b) $f(x) = e^{-\|x\|^2}$
   (c) $f(x) = x_1^2 - 2x_1x_2 + \frac{1}{3}x_2^3 - 8x_2$
   (d) $f(x) = (2x_1 - x_2)^2 + (x_2 - x_3)^2 + (x_3 - 1)^2$
   (e) $f(x_1, x_2) := x_1^2 + 2x_2x_2 + x_2^2 - 8x_1 - 8x_2$
   (f) $f(x_1, x_2) = \ln(e^{x_1} + e^{x_2})$, where $\ln(\mu)$ is the natural log of μ for $\mu > 0$ and is $+\infty$ for $\mu \le 0$.

2. Suppose $f : \mathbb{R}^n \to \mathbb{R}$ is a differentiable convex function and let $\Omega \subset \mathbb{R}^n$ is a closed convex set. Use the subdifferential inequality to show that $\bar{x}$ is a global solution to the problem $\min_{x \in \Omega} f(x)$ if and only if $f'(\bar{x}; x - \bar{x}) \ge 0$ for all $x \in \Omega$.

3. Compute the Lagrangian dual to each of the following problems.

   (a) $\min\left\{\frac{1}{2}\|x\|_2^2 \;\middle|\; a^T x \le \alpha\right\}$ where $a \in \mathbb{R}^n \setminus \{0\}$ and $\alpha \in \mathbb{R}$.
   (b) $\min\left\{\frac{1}{2}\|x\|_2^2 \;\middle|\; a^T x = \alpha,\ b^T x \le \beta\right\}$ where $a, b \in \mathbb{R}^n \setminus \{0\}$ with a and b linearly independent and $\alpha, \beta \in \mathbb{R}$.
   (c) $\min\left\{\frac{1}{2}\|x\|_2^2 \;\middle|\; a^T x \le \alpha,\ b^T x \le \beta\right\}$ where $a, b \in \mathbb{R}^n \setminus \{0\}$ with a and b linearly independent and $\alpha, \beta \in \mathbb{R}$.
   (d) $\min\left\{\frac{1}{2}x^T H x \;\middle|\; \mathsf{e}^T x = 1 \quad \text{and} \quad r^T x = \mu\right\}$ where $H \in \mathbb{S}^n_{++}$, $\mathsf{e} := (1, 1, \ldots, 1)^T \in \mathbb{R}^n$ is the vector of all ones, e and $r \in \mathbb{R}^n$ are linearly independent, and $\mu \in \mathbb{R}$.
   (e) $\min\left\{\frac{1}{2}x^T H x \;\middle|\; \mathsf{e}^T x = 1 \quad \text{and} \quad r^T x \ge \mu\right\}$ where $H \in \mathcal{S}^n_{++}$, $\mathsf{e} := (1, 1, \ldots, 1)^T \in \mathbb{R}^n$ is the vector of all ones, e and $r \in \mathbb{R}^n$ are linearly independent, and $\mu \in \mathbb{R}$.
   (f) $\min\left\{4x_1^2 + 4x_1x_2 + x_2^2 - 8x_1 - 4x_2 \;\middle|\; x_1 + x_2 \le 0,\ x_1 - x_2 \le 0\right\}$
   (g) Let $Q \in \mathbb{S}^n_{++}$, $c \in \mathbb{R}^n$, $A \in \mathbb{R}^{m \times n}$, and $b \in \mathbb{R}^m$ and compute the Lagrangian dual to the problem

$$\begin{array}{lll} \mathcal{P} & \text{minimize} & \frac{1}{2}x^T Q x + c^T x \\ & \text{subject to} & Ax \le b,\ 0 \le x\ . \end{array}$$

4. Let $A \in \mathbb{R}^{m \times n}$ and $b \in \mathbb{R}^m$. Consider the optimization problem

$$\begin{array}{lll} \mathcal{P} & \text{minimize} & \frac{1}{2}\|Ax - b\|_2^2 \\ & \text{subject to} & \|x\|_1 \le 1\ . \end{array}$$

(a) Show that this problem is equivalent to the problem

$$\hat{\mathcal{P}} \qquad \begin{array}{ll} \text{minimize}_{(x,z,w)} & \frac{1}{2}\|w\|_2^2 \\ \text{subject to} & Ax - b = w, \\ & -z \le x \le z \ \ \text{and} \ \ e^T z \le 1, \end{array}$$

where e is the vector of all ones.

(b) What is the Lagrangian for $\hat{\mathcal{P}}$?

(c) Show that the Lagrangian dual for $\hat{\mathcal{P}}$ is the problem

$$\mathcal{D} \qquad \max -\frac{1}{2}\|y\|_2^2 - y^T b - \|A^T y\|_\infty \qquad = \qquad -\min \frac{1}{2}\|y-b\|_2^2 + \|A^T y\|_\infty - \frac{1}{2}\|b\|_2^2 \,.$$

This is also the Lagrangian dual for $\mathcal{P}$.

5. Consider the functions

$$f(x) = \frac{1}{2}x^T Q x - c^T x$$

and

$$f_t(x) = \frac{1}{2}x^T Q x - c^T x + t\phi(x),$$

where $t > 0$, $Q \in \mathbb{S}^n_+$, $c \in \mathbb{R}^n$, and $\phi : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ is given by

$$\phi(x) = \begin{cases} -\sum_{i=1}^n \ln x_i & \text{, if } x_i > 0, \ i = 1, 2, \ldots, n, \\ +\infty & \text{, otherwise.} \end{cases}$$

(a) Show that ϕ is a convex function.

(b) Show that both f and f_t are convex functions.

(c) Show that the solution to the problem $\min f_t(x)$ always exists and is unique.

(d) Let $\{t_i\}$ be a decreasing sequence of positive real scalars with $t_i \downarrow 0$, and let x^i be the solution to the problem $\min f_{t_i}(x)$. Show that if the sequence $\{x^i\}$ has a cluster point $\bar{x}$, then $\bar{x}$ must be a solution to the problem $\min\{f(x) : 0 \le x\}$.

Hint: Use the KKT conditions for the QP $\min\{f(x) : 0 \le x\}$.