University of Waterloo
Department of Electrical and Computer Engineering
Winter, 2019

ECE 602: Introduction to Optimization

FINAL EXAMINATION

Surname								
Legal Given Name(s)								
UW Student ID Number								

Instruction:

1. There are 100 points in total.
2. This is a written, open-book exam (only annotated lecture slides are allowed). Please turn off all electronic media and store them under your desk.
3. Be neat. Poor presentation will be penalized.
4. **No questions will be answered during the exam**. If in doubt, state your assumption(s) and continue.
5. Do not leave during the examination period without permission.
6. Do not stand up until all the exams have been picked up.

Do well!

Question 1 (20 points)

Suppose $p < 1$, $p \neq 0$. Show that the function

$$f(x) = \left(\sum_{i=1}^{n} x_i^p \right)^{1/p}, \quad \text{with} \quad \textbf{dom} f = \mathbb{R}_{++}^n,$$

is concave.

<u>Hint:</u> To prove that $\nabla^2 f(x)$ is negative semi-definite, it is sufficient to show that

$$v^T \nabla^2 f(x) v \leq 0, \quad \forall v \in \mathbb{R}^n.$$

Question 2 (20 points)

Consider the problem, with variable $x \in \mathbb{R}^n$,

$$\begin{array}{ll} \text{minimize} & c^T x \\ \text{subject to} & Ax \preceq b \quad \text{for all } \; A \in \mathcal{A}, \end{array}$$

where $\mathcal{A} \in \mathbb{R}^{m \times n}$ is the set

$$\mathcal{A} = \left\{ A \in \mathbb{R}^{m \times n} \mid \bar{A}_{ij} - V_{ij} \leq A_{ij} \leq \bar{A}_{ij} + V_{ij}, \; i = 1, \ldots, m, \; j = 1, \ldots, n \right\},$$

with some given matrices $\bar{A}$ and V. This problem can be interpreted as an LP where each coefficient of A is only known to lie in an interval, and we require that x must satisfy the constraints for all possible values of the coefficients.

Express this problem as an LP.

Question 3 (20 points)

Consider the following unconstrained geometric program with the variable $x \in \mathbb{R}^n$

$$\text{minimize}_x \quad \log \left(\sum_{i=1}^{m} \exp(a_i^T x + b_i) \right),$$

for some given $a_1, a_2, \ldots, a_m \in \mathbb{R}^n$ and $b_1, b_2, \ldots, b_m \in \mathbb{R}$. This problem can be reformulated in an equivalent constrained form by introducing an additional optimization variable $y \in \mathbb{R}^m$ defined as

$$y = Ax + b,$$

with $A = [a_1 \, a_2 \, \ldots \, a_m]^T \in \mathbb{R}^{m \times n}$ and $b = [b_1 \, b_2 \, \ldots \, b_m]^T \in \mathbb{R}^m$. The equivalent problem can then be defined as

$$\begin{array}{ll} \text{minimize}_{(x,y)} & f_0(y) = \log \left(\sum_{i=1}^{m} \exp y_i \right) \\ \text{subject to} & Ax + b = y. \end{array}$$

Derive the dual form of the above optimization problem.

Question 4 (20 points)

Consider the following linear program

$$\begin{aligned} &\text{minimize } c^T x \\ &\text{subject to } Ax = b \\ &\qquad\qquad l \preceq x \preceq u \end{aligned}$$

with some given $c, u, l \in \mathbb{R}^n$, $b \in \mathbb{R}^m$, and $A \in \mathbb{R}^{m \times n}$. This problem can be equivalently reformulated as

$$\begin{aligned} &\text{minimize } f_0(x) \\ &\text{subject to } Ax = b, \end{aligned}$$

for some extended-value function $f_0 : \mathbb{R}^n \to (-\infty, +\infty]$.

a) Find an explicit definition of the function $f_0(x)$. Is the function convex? Is it closed?

b) Find the dual function $g(\nu)$ for the equivalent problem above, where $\nu \in \mathbb{R}^m$ is the vector of Lagrange multipliers.

c) Formulate the corresponding dual problem.

Hint: The final answer should involve maximization (with respect to $\nu \in \mathbb{R}^m$) of a piece-wise linear function whose definition depends on the same given c, b, u, l and A, as well as on the use of functions $(\cdot)^+ : \mathbb{R}^n \to \mathbb{R}^n_+$ and $(\cdot)^- : \mathbb{R}^n \to \mathbb{R}^n_+$ and $(\cdot)^- : \mathbb{R}^n \to \mathbb{R}^n_+$ defined as given by

$$(y)_i^+ = \max\{y_i, 0\} \qquad \text{and} \qquad (y)_i^- = \max\{-y_i, 0\}, \qquad \forall y \in \mathbb{R}^n.$$

Question 5 (20 points)

Formulate the following robust approximation problems as LPs, QP, SOCPs, or SDPs. For each subproblem, consider the ℓ_1-, ℓ_2-, and the ℓ_∞-norms.

a) *Stochastic robust approximation with a finite set of parameter values:*

$$\text{minimize } \sum_{i=1}^{k} p_i \|A_i x - b\|,$$

where $p \succeq 0$ and $\mathbf{1}^T p = 1$. Both $A_1, A_2, \ldots, A_k \in \mathbb{R}^{m \times n}$ and $b \in \mathbb{R}^m$ are assumed to be given.

b) *Worst-case robust approximation with coefficient bounds:*

$$\text{minimize } \sup_{A \in \mathcal{A}} \|Ax - b\|,$$

where

$$\mathcal{A} = \left\{ A \in \mathbb{R}^{m \times n} \mid l_{ij} \le a_{ij} \le u_{ij},\ i = 1, \ldots, m,\ j = 1, \ldots, n \right\}.$$

Note that here the uncertainty set is described by giving upper and lower bounds for the components of A. It is assumed that $\{l_{ij}\}$ and $\{u_{ij}\}$ are given and that $l_{ij} < u_{ij}$, $\forall (i, j)$.

SOLUTIONS

Question 1

Solution. The first derivatives of f are given by

$$\frac{\partial f(x)}{\partial x_i} = (\sum_{i=1}^{n} x_i^p)^{(1-p)/p} x_i^{p-1} = \left(\frac{f(x)}{x_i}\right)^{1-p}.$$

The second derivatives are

$$\frac{\partial^2 f(x)}{\partial x_i \partial x_j} = \frac{1-p}{x_i}\left(\frac{f(x)}{x_i}\right)^{-p}\left(\frac{f(x)}{x_j}\right)^{1-p} = \frac{1-p}{f(x)}\left(\frac{f(x)^2}{x_i x_j}\right)^{1-p}$$

for $i \neq j$, and

$$\frac{\partial^2 f(x)}{\partial x_i^2} = \frac{1-p}{f(x)}\left(\frac{f(x)^2}{x_i^2}\right)^{1-p} - \frac{1-p}{x_i}\left(\frac{f(x)}{x_i}\right)^{1-p}.$$

We need to show that

$$y^T \nabla^2 f(x) y = \frac{1-p}{f(x)}\left(\left(\sum_{i=1}^{n} \frac{y_i f(x)^{1-p}}{x_i^{1-p}}\right)^2 - \sum_{i=1}^{n} \frac{y_i^2 f(x)^{2-p}}{x_i^{2-p}}\right) \leq 0$$

This follows by applying the Cauchy-Schwarz inequality $a^T b \leq \|a\|_2 \|b\|_2$ with

$$a_i = \left(\frac{f(x)}{x_i}\right)^{-p/2}, \qquad b_i = y_i \left(\frac{f(x)}{x_i}\right)^{1-p/2},$$

and noting that $\sum_i a_i^2 = 1$.

Question 2

Solution. The problem is equivalent to

$$\begin{array}{ll} \text{minimize} & c^T x \\ \text{subject to} & \bar{A}x + V|x| \preceq b \end{array}$$

where $|x| = (|x_1|, |x_2|, \ldots, |x_n|)$. This in turn is equivalent to the LP

$$\begin{array}{ll} \text{minimize} & c^T x \\ \text{subject to} & \bar{A}x + Vy \preceq b \\ & -y \preceq x \preceq y \end{array}$$

with variables $x \in \mathbf{R}^n$, $y \in \mathbf{R}^n$.

Question 3

Now let us reformulate the problem as

$$\begin{array}{ll} \text{minimize} & f_0(y) \\ \text{subject to} & Ax + b = y. \end{array}$$

Here we have introduced new variables y, as well as new equality constraints $Ax + b = y$.

The Lagrangian of the reformulated problem is

$$L(x, y, \nu) = f_0(y) + \nu^T(Ax + b - y).$$

To find the dual function we minimize L over x and y. Minimizing over x we find that $g(\nu) = -\infty$ unless $A^T\nu = 0$, in which case we are left with

$$g(\nu) = b^T\nu + \inf_y (f_0(y) - \nu^T y) = b^T\nu - f_0^*(\nu),$$

where f_0^* is the conjugate of f_0. The dual problem can therefore be expressed as

$$\begin{array}{ll} \text{maximize} & b^T\nu - f_0^*(\nu) \\ \text{subject to} & A^T\nu = 0. \end{array}$$

The conjugate of the log-sum-exp function is

$$f_0^*(\nu) = \begin{cases} \sum_{i=1}^m \nu_i \log \nu_i & \nu \succeq 0,\ \mathbf{1}^T\nu = 1 \\ \infty & \text{otherwise} \end{cases}$$

so the dual of the reformulated problem can be expressed as

$$\begin{array}{ll} \text{maximize} & b^T\nu - \sum_{i=1}^m \nu_i \log \nu_i \\ \text{subject to} & \mathbf{1}^T\nu = 1 \\ & A^T\nu = 0 \\ & \nu \succeq 0, \end{array}$$

which is an entropy maximization problem.

Question 4

We can, of course, derive the dual of this linear program. The dual will have a Lagrange multiplier ν associated with the equality constraint, λ_1 associated with the inequality constraint $x \preceq u$, and λ_2 associated with the inequality constraint $l \preceq x$. The dual is

$$\begin{array}{ll} \text{maximize} & -b^T\nu - \lambda_1^T u + \lambda_2^T l \\ \text{subject to} & A^T\nu + \lambda_1 - \lambda_2 + c = 0 \\ & \lambda_1 \succeq 0, \quad \lambda_2 \succeq 0. \end{array}$$

Instead, let us first reformulate the problem as

$$\begin{array}{ll} \text{minimize} & f_0(x) \\ \text{subject to} & Ax = b, \end{array}$$

where we define

$$f_0(x) = \begin{cases} c^T x & l \preceq x \preceq u \\ \infty & \text{otherwise.} \end{cases}$$

The dual function for the problem is

$$\begin{aligned} g(\nu) &= \inf_{l \preceq x \preceq u} \left(c^T x + \nu^T (Ax - b) \right) \\ &= -b^T \nu - u^T (A^T \nu + c)^- + l^T (A^T \nu + c)^+ \end{aligned}$$

where $y_i^+ = \max\{y_i, 0\}$, $y_i^- = \max\{-y_i, 0\}$. So here we are able to derive an analytical formula for g, which is a concave piecewise-linear function.

The dual problem is the unconstrained problem

$$\text{maximize} \quad -b^T \nu - u^T (A^T \nu + c)^- + l^T (A^T \nu + c)^+$$

Question 5

(a) *Stochastic robust approximation with a finite set of parameter values, i.e.*, the sum-of-norms problem

$$\text{minimize} \quad \sum_{i=1}^k p_i \|A_i x - b\|$$

where $p \succeq 0$ and $\mathbf{1}^T p = 1$.

Solution.

- ℓ_1-norm:

$$\begin{array}{ll} \text{minimize} & \sum_{i=1}^k p_i \mathbf{1}^T y_i \\ \text{subject to} & -y_i \preceq A_i x - b \preceq y_i, \quad i = 1, \ldots, k. \end{array}$$

  An LP with variables $x \in \mathbf{R}^n$, $y_i \in \mathbf{R}^m$, $i = 1, \ldots, k$.

- ℓ_2-norm:

$$\begin{array}{ll} \text{minimize} & p^T y \\ \text{subject to} & \|A_i x - b\|_2 \leq y_i, \quad i = 1, \ldots, k. \end{array}$$

  An SOCP with variables $x \in \mathbf{R}^n$, $y \in \mathbf{R}^k$.

- ℓ_∞-norm:

$$\begin{array}{ll} \text{minimize} & p^T y \\ \text{subject to} & -y_i \mathbf{1} \preceq A_i x - b \leq y_i \mathbf{1}, \quad i = 1, \ldots, k. \end{array}$$

  An LP with variables $x \in \mathbf{R}^n$, $y \in \mathbf{R}^k$.

(b) *Worst-case robust approximation with coefficient bounds:*

$$\text{minimize} \quad \sup_{A \in \mathcal{A}} \|Ax - b\|$$

where

$$\mathcal{A} = \{A \in \mathbf{R}^{m \times n} \mid l_{ij} \le a_{ij} \le u_{ij},\ i = 1, \ldots, m,\ j = 1, \ldots, n\}.$$

Here the uncertainty set is described by giving upper and lower bounds for the components of A. We assume $l_{ij} < u_{ij}$.

Solution. We first note that

$$\begin{aligned} \sup_{l_{ij} \le a_{ij} \le u_{ij}} |a_i^T x - b_i| &= \sup_{l_{ij} \le a_{ij} \le u_{ij}} \max\{a_i^T x - b_i, -a_i^T x + b_i\} \\ &= \max\{\sup_{l_{ij} \le a_{ij} \le u_{ij}} (a_i^T x - b_i), \sup_{l_{ij} \le a_{ij} \le u_{ij}} (-a_i^T x + b_i)\}. \end{aligned}$$

Now,

$$\sup_{l_{ij} \le a_{ij} \le u_{ij}} (\sum_{j=1}^n a_{ij} x_j - b_i) = \bar{a}_i^T x - b_i + \sum_{j=1}^n v_{ij} |x_j|$$

where $\bar{a}_{ij} = (l_{ij} + u_{ij})/2$, and $v_{ij} = (u_{ij} - l_{ij})/2$, and

$$\sup_{l_{ij} \le a_{ij} \le u_{ij}} (-\sum_{j=1}^n a_{ij} x_j + b_i) = -\bar{a}_i^T x + b_i + \sum_{j=1}^n v_{ij} |x_j|.$$

Therefore

$$\sup_{l_{ij} \le a_{ij} \le u_{ij}} |a_i^T x - b_i| = |\bar{a}_i^T x - b_i| + \sum_{j=1}^n v_{ij} |x_j|.$$

- ℓ_1-norm:

  $$\text{minimize} \quad \textstyle\sum_{i=1}^m \left(|\bar{a}_i^T x - b_i| + \sum_{j=1}^n v_{ij} |x_j| \right).$$

  This can be expressed as an LP

  $$\begin{array}{ll} \text{minimize} & \mathbf{1}^T (y + Vw) \\ & -y \preceq \bar{A}x - b \preceq y \\ & -w \preceq x \preceq w. \end{array}$$

  The variables are $x \in \mathbf{R}^n$, $y \in \mathbf{R}^m$, $w \in \mathbf{R}^n$.

- ℓ_2-norm:

$$\text{minimize} \quad \sum_{i=1}^m \left(|\bar{a}_i^T x - b_i| + \sum_{j=1}^n v_{ij}|x_j| \right)^2.$$

  This can be expressed as an SOCP

$$\begin{array}{ll} \text{minimize} & t \\ \text{subject to} & -y \preceq \bar{A}x - b \preceq y \\ & -w \preceq x \preceq w \\ & \|y + Vw\|_2 \leq t. \end{array}$$

  The variables are $x \in \mathbf{R}^n$, $y \in \mathbf{R}^m$, $w \in \mathbf{R}^n$, $t \in \mathbf{R}$.

- ℓ_∞-norm:

$$\text{minimize} \quad \max_{i=1,\ldots,m} \left(|\bar{a}_i^T x - b_i| + \sum_{j=1}^n v_{ij}|x_j| \right).$$

  This can be expressed as an LP

$$\begin{array}{ll} \text{minimize} & t \\ & -y \preceq \bar{A}x - b \preceq y \\ & -w \preceq x \preceq w \\ & -t\mathbf{1} \preceq y + Vw \leq t\mathbf{1}. \end{array}$$

  The variables are $x \in \mathbf{R}^n$, $y \in \mathbf{R}^m$, $w \in \mathbf{R}^n$, $t \in \mathbf{R}$.