Mathematical Optimization

Exam April 18, 2023, 8:45 - 11:45

No additional materials may be used during this exam (no notes, calculators, etc.). With this exam a list of theorems and lemmata is provided. In your proofs, you may use definitions from the lecture notes and the theorems and lemmata from the list without providing a proof (reference the theorem/lemma that you use). In addition, you may use all results from Appendix A and all theorems, lemmata, corollaries and propositions from Chapters 6 (Convex Sets), 7 (Convex Functions) and 9 (Iterative Optimization Methods) in the Lecture Notes (v. January 24, 2023) with a reference like "We know that...".

This exam has 8 exercises.

1. Given
$$\mathbf{A} = \begin{pmatrix} 1 & 0 & 0 \\ 7 & 1 & 0 \\ -3 & 1 & 1 \end{pmatrix} \cdot \begin{pmatrix} -14 & 8 & 12 \\ 0 & 32 & 49 \\ 0 & 0 & 0 \end{pmatrix} \qquad \text{and} \qquad \mathbf{b} = \begin{pmatrix} -1 \\ 2 \\ 3 \end{pmatrix},$$
either give a solution to $\mathbf{Ax} = \mathbf{b}$ or use Corollary 3 to prove that no such solution exists.

2. Recall that $L(\mathbf{v}_1, \ldots, \mathbf{v}_n) = \{\sum_{j=1}^n \mathbf{v}_j \lambda_j \mid \lambda_j \in \mathbb{Z}\}$ denotes the lattice generated by the vectors $\mathbf{v}_1, \ldots, \mathbf{v}_n$.

   (a) Compute linearly independent vectors $\mathbf{c}_1$ and $\mathbf{c}_2$, such that
   $$L(\mathbf{c}_1, \mathbf{c}_2) = L(\begin{pmatrix} 2 \\ 6 \end{pmatrix}, \begin{pmatrix} 3 \\ 1 \end{pmatrix}, \begin{pmatrix} 12 \\ 12 \end{pmatrix}).$$

   (b) Decide if there is an integer solution $\mathbf{x}$ to the system
   $$\begin{pmatrix} 2 & 3 & 12 \\ 6 & 1 & 12 \end{pmatrix} \mathbf{x} = \begin{pmatrix} 2 \\ -10 \end{pmatrix}.$$
   If such $\mathbf{x}$ exists, provide one. If no such $\mathbf{x}$ exists, provide a vector $\mathbf{y}$ that proofs it.

3. Use the Fourier-Motzkin procedure to show that all solutions to the system:
$$\begin{array}{rcrcrcr} 2x & - & y & + & 6z & \leq & 2 \\ -3x & + & 3y & - & 6z & \leq & 15 \\ x & + & 2y & - & z & \leq & 13 \\ x & - & 3y & & & \leq & -17 \end{array}$$
satisfy $0 \leq y \leq 8$.

4. Let W be a linear subspace of $\mathbb{R}^n$. Consider the problem of finding a best approximation $\hat{\mathbf{x}} \in W$ of $\mathbf{x} \in \mathbb{R}^n$. That is, finding $\hat{\mathbf{x}} \in W$ such that

$$\|\mathbf{x} - \hat{\mathbf{x}}\|_2 = \min_{\mathbf{y} \in W} \|\mathbf{x} - \mathbf{y}\|_2 .$$

Suppose we found $\mathbf{y} \in W$ such that $\mathbf{x} - \mathbf{y}$ is orthogonal to every $\mathbf{w} \in W$.
Prove that $\mathbf{y}$ is the unique best approximation of $\mathbf{x}$ in $\mathbb{R}^n$.

5. Prove that a solution to $\mathbf{A}^T\mathbf{y} = \mathbf{c}$, $\mathbf{y} \geq \mathbf{0}$ exists if and only if $\mathbf{A}\mathbf{x} \leq \mathbf{0}$ implies $\mathbf{c}^T\mathbf{x} \leq \mathbf{0}$.

6. (a) Show that $f(\mathbf{x}) = \|x\|$ defines a convex function $f : \mathbb{R}^n \to \mathbb{R}$. Here, $\|x\|$ denotes any norm on $\mathbb{R}^n$.
   (b) Let $g : \mathbb{R}^n \to I, I \subseteq \mathbb{R}$ be convex and $f : I \to \mathbb{R}$ be convex and non-decreasing. Show that the composition $f \circ g(\mathbf{x}) = f(g(\mathbf{x}))$ of f and g is convex.
   (c) Consider the function $f(x, y) = e^{x-y}$. Prove that

$$f(x, y) \geq 1 + x - y \quad \text{for all } (x, y) \in \mathbb{R}^2.$$

7. Let $f(\mathbf{x}) = \frac{1}{2}x_1^4 + 2x_1x_2 + 2x_1 + (1 + x_2)^2$.
   (a) Determine the critical points and the local minimizer(s) of f.
   (b) Does f have (a) global minimizer(s)? Motivate!

8. (a) Pick a suitable series of step sizes t_k (argue why your pick is suitable) and apply two steps of the subgradient method to

$$f(\mathbf{x}) = \max\{x_1 + x_2, -x_1 + x_2, x_1 - x_2, -x_1 - x_2\}$$

   starting at $\mathbf{x}_0 = (0, \frac{1}{\sqrt{2}})^T$.
   (b) Let $f : \mathbb{R}^2 \to \mathbb{R}$ be defined by $f(x, y) = \frac{1}{3}x^3 + 2xy + x + y^2$. Apply one step of Newton's method to f, starting from $(x_0, y_0) = (2, 0)$. Is the direction a descent direction?

Points: 90 + 10 = 100

1.		:	8 pt.	6.	(a)	:	6 pt.
2.	(a)	:	8 pt.		(b)	:	6 pt.
	(b)	:	5 pt.		(c)	:	6 pt.
3.		:	6 pt.	7.	(a)	:	8 pt.
4.		:	8 pt.		(b)	:	5 pt.
5.		:	8 pt.	8.	(a)	:	8 pt.
					(b)	:	8 pt.

Script Mathematical Optimization (2022/2023)

The following results will be provided during the examinations of Mathematical Optimization, and can be used without proof, referring to the name.

Lemma 1 (Taylor's formula)**.** For a C^2-function $f : U \to \mathbb{R}$, with $U \subseteq \mathbb{R}^n$, and $\|\mathbf{d}\|$ small enough:

$$f(\mathbf{x}_0 + \mathbf{d}) = f(\mathbf{x}_0) + \nabla f(\mathbf{x}_0)\mathbf{d} + \frac{1}{2}\mathbf{d}^T \nabla^2 f(\mathbf{x}_0)\mathbf{d} + o(\|\mathbf{d}\|^2)$$

or for, additionally, some $\tau \in (0, 1)$:

$$f(\mathbf{x}_0 + \mathbf{d}) = f(\mathbf{x}_0) + \nabla f(\mathbf{x}_0)\mathbf{d} + \frac{1}{2}\mathbf{d}^T \nabla^2 f(\mathbf{x}_0 + \tau\mathbf{d})\mathbf{d}\,.$$

Corollary 2 (LU-factorization)**.** For every matrix $\mathbf{A} \in \mathbb{R}^{m\times n}$, there exists an $(m \times m)$-permutation matrix $\mathbf{P}$ and an invertible lower triangular matrix $\mathbf{M} \in \mathbb{R}^{m\times m}$ such that $\mathbf{U} = \mathbf{MPA}$ is upper triangular.

Corollary 3 (Gale's Theorem)**.** Let $\mathbf{A} \in \mathbb{R}^{m\times n}$ and $\mathbf{b} \in \mathbb{R}^m$ be given.
Then exactly one of the following alternatives is true:

(I) There exists some $\mathbf{x} \in \mathbb{R}^n$ such that $\mathbf{Ax} = \mathbf{b}$.

(II) There exists a vector $\mathbf{y} \in \mathbb{R}^m$ such that $\mathbf{y}^T\mathbf{A} = \mathbf{0}^T$ and $\mathbf{y}^T\mathbf{b} \neq 0$.

Theorem 4 (Integer solutions to linear system of equations)**.** Let $\mathbf{A} \in \mathbb{Z}^{m\times n}$ and $\mathbf{b} \in \mathbb{Z}^m$ be given. Then exactly one of the following statements is true:

(I) There exists some $\mathbf{x} \in \mathbb{Z}^n$ such that $\mathbf{Ax} = \mathbf{b}$.

(II) There exists some $\mathbf{y} \in \mathbb{R}^m$ such that $\mathbf{y}^T\mathbf{A} \in \mathbb{Z}^n$ and $\mathbf{y}^T\mathbf{b} \notin \mathbb{Z}$.

Corollary 5 (Identification of positive (semi-)definite matrices)**.** Let $\mathbf{A} \in \mathbb{R}^{n\times n}$ be a symmetric matrix and $\mathbf{Q} \in \mathbb{R}^{n\times n}$ an invertible matrix such that $\mathbf{D} = \mathbf{QAQ}^T$ is diagonal. Then

(a) $\mathbf{A}$ is *positive semidefinite* if and only if all diagonal elements of $\mathbf{D}$ are non-negative.

(b) $\mathbf{A}$ is *positive definite* if and only if all diagonal elements of $\mathbf{D}$ are strictly positive.

Corollary 6 (Identification of 2×2 positive (semi-)definite matrices). Let $\mathbf{A} \in \mathbb{R}^{2\times 2}$ be a symmetric matrix. Then

(a) $\mathbf{A}$ is *positive semidefinite* if and only if all diagonal elements and the determinant of $\mathbf{A}$ are non-negative.

(b) $\mathbf{A}$ is *positive definite* if and only if all diagonal elements and the determinant of $\mathbf{A}$ are strictly positive.

Theorem 7 (Spectral Theorem for Symmetric Matrices). Let $\mathbf{A} \in \mathbb{R}^{n\times n}$ be a symmetric matrix. Then there exists a matrix $\mathbf{Q} \in \mathbb{R}^{n\times n}$ and eigenvalues $\lambda_1, \dots, \lambda_n$ of $\mathbf{A}$ such that

$$\mathbf{Q}^T\mathbf{Q} = \mathbf{I} \quad \text{and} \quad \mathbf{Q}^T\mathbf{A}\mathbf{Q} = \operatorname{diag}(\lambda_1, \dots, \lambda_n) .$$

Note: the columns of $\mathbf{Q}$ form an orthonormal basis of $\mathbb{R}^n$, consisting of eigenvectors of $\mathbf{A}$.

Theorem 8 (Farkas Lemma). Let $\mathbf{A} \in \mathbb{R}^{m\times n}$ and $\mathbf{b} \in \mathbb{R}^m$ be given.
Then exactly one of the following alternatives is true:

(I) $\mathbf{Ax} \le \mathbf{b}$ is feasible.

(II) There exists a vector $\mathbf{y} \ge \mathbf{0}$ such that $\mathbf{y}^T\mathbf{A} = \mathbf{0}^T$ and $\mathbf{y}^T\mathbf{b} < 0$.

Theorem 9 (Strong Duality). Let $\mathbf{A} \in \mathbb{R}^{m\times n}$, $\mathbf{b} \in \mathbb{R}^m$, $\mathbf{c} \in \mathbb{R}^n$ and be given, and suppose that either:

(I) there exists some $\mathbf{x} \in \mathbb{R}^n$ such that $\mathbf{Ax} \le \mathbf{b}$

or

(II) there exists some $\mathbf{y} \in \mathbb{R}^m$ such that $\mathbf{A}^T\mathbf{y} = \mathbf{c}$ and $\mathbf{y} \ge \mathbf{0}$.

Then

$$\max\{\mathbf{c}^T\mathbf{x} \,|\, \mathbf{Ax} \le \mathbf{b}\} \;=\; \min\{\mathbf{b}^T\mathbf{y} \,|\, \mathbf{A}^T\mathbf{y} = \mathbf{c}, \mathbf{y} \ge \mathbf{0}\} .$$

If both (I) and (II) are feasible, then optimal solutions $\mathbf{x}$ of (I) and $\mathbf{y}$ of (II) exist and satisfy $\mathbf{c}^T\mathbf{x} = \mathbf{b}^T\mathbf{y}$.

Lemma 10 (Necessary optimality conditions). Let f be a C^2-function on $\mathbb{R}^n$. Then each local minimizer $\overline{\boldsymbol{x}} \in \mathbb{R}^n$ of f satisfies:

$$\begin{array}{lllll} \text{(a)} & \text{(First order condition)} & \nabla f(\overline{\boldsymbol{x}}) & = & \mathbf{0}^T \\ \text{(b)} & \text{(Second order condition)} & \mathbf{d}^T\nabla^2 f(\overline{\boldsymbol{x}})\mathbf{d} & \ge & 0 \quad \text{for all } \mathbf{d} \in \mathbb{R}^n . \end{array}$$

Lemma 11 (Sufficient optimality condition). Let f be a C^2-function on $\mathbb{R}^n$ and $\overline{\boldsymbol{x}} \in \mathbb{R}^n$ such that $\nabla f(\overline{\boldsymbol{x}}) = \mathbf{0}^T$ holds. Then $\overline{\boldsymbol{x}}$ is a strict local minimizer of f, provided $\overline{\boldsymbol{x}}$ satisfies

$$\mathbf{d}^T\nabla^2 f(\overline{\boldsymbol{x}})\mathbf{d} > 0 \qquad \text{for all } \mathbf{d} \in \mathbb{R}^n \setminus \{\mathbf{0}\}.$$