EECS 127: Optimization Models in Engineering **Spring 2025**

Convexity on a Line, Convex Duality, Farkas' Lemma

1 Review: Convexity on a Line

Last Time: Convex Sets & Functions. Recall some characterizations of convexity of functions, which we can use

1. First-order convexity: $\forall x, y \in \text{dom}(f), f(x) \geq f(y) + \nabla f(y)^T(x - y)$
2. Second-order convexity: $\forall x \in \text{dom}(f), \nabla^2 f(x) \geq 0$
3. Convexity along a line (Today)

We now detail this procedure (#3) to check for convexity on a given function f along a line.

Theorem 1.1. *Pick starting point x_0 and a direction $\vec{z}$. To show that f is convex along this line given by $\vec{z}$, we want to check that for all sufficiently small $t \geq 0$[1], the function $f(x_0 + t\vec{z})$ is convex as a function of t.*

Note that if we show convexity for all x_0 and direction z, we have shown that f is convex. We do this in our following example.

Example 1.2. Consider the function $\mathbf{X} \mapsto \log\det(\mathbf{X}^{-1})$, which is convex on $\mathbf{X} > 0$.
We want to check that $\mathbf{X}_0 + t\mathbf{Z}$ is convex for all $\mathbf{X} > 0$ to conclude that the above function is convex. To do so, we must be sure that $\mathbf{X}_0 + t\mathbf{Z}$ is positive semi-definite (so as for the $\log(\cdot)$ function to operate on a well-defined domain). To ensure this, we impose the constraint that $\mathbf{Z}$ is symmetric. We now verify with matrix algebra that log $\det((\mathbf{X}_0 + t\mathbf{Z})^{-1})$ is convex as a function of t:

$$\log\det\left((\mathbf{X}_0 + t\mathbf{Z})^{-1}\right) = -\log\det(\mathbf{X}_0 + t\mathbf{Z}) = -\log\det\left(\mathbf{X}_0^{\frac{1}{2}}\left(\mathbf{I} + t\mathbf{X}_0^{-\frac{1}{2}}\mathbf{Z}\mathbf{X}_0^{-\frac{1}{2}}\right)\mathbf{X}_0^{\frac{1}{2}}\right)$$

$$= -\log\det(\mathbf{X}_0) - \log\det\left(\mathbf{I} + \mathbf{X}_0^{-\frac{1}{2}}\mathbf{Z}\mathbf{X}_0^{-\frac{1}{2}}\right)$$

For clarity, we WTS that this final expression is convex for this arbitrarily chosen (with constraints) $\mathbf{X}_0$ and $\mathbf{Z}$ to show that $\mathbf{X} \mapsto \log\det(\mathbf{X}^{-1})$ is convex. We approach this with a property of the determinant: the determinant of a matrix is equal to the product of its eigenvalues $\implies$ log-determinant of a matrix, in our case $\mathbf{I} + \mathbf{X}_0^{-\frac{1}{2}}\mathbf{Z}\mathbf{X}_0^{-\frac{1}{2}}$, is equal to the sum of its eigenvalues. The eigenvalues of the identity matrix are all equal to one and so we can re-express this function as:

$$-\log\det(\mathbf{X}_0) - \sum_{i=1}^{n} \log\left(1 + t\lambda_i\left(\mathbf{X}_0^{-\frac{1}{2}}\mathbf{Z}\mathbf{X}_0^{-\frac{1}{2}}\right)\right)$$

[1]The value of t must be sufficiently small because we must have that $x_0 + t\vec{z} \in \text{dom}(f)$

where $\lambda_i\left(\mathbf{X}_0^{-\frac{1}{2}}\mathbf{Z}\mathbf{X}_0^{-\frac{1}{2}}\right)$ represents the i-th eigenvalue of $\mathbf{X}_0^{-\frac{1}{2}}\mathbf{Z}\mathbf{X}_0^{-\frac{1}{2}}$.
First, recall that the function $-\log(x)$ is trivially convex for $\mathbb{R}^+$. Consider this term of interest:

$$-\sum_{i=1}^{n}\log\left(1+t\lambda_i\left(\mathbf{X}_0^{-\frac{1}{2}}\mathbf{Z}\mathbf{X}_0^{-\frac{1}{2}}\right)\right).$$

Note that each term $-\log(1+t\lambda_i)$ is convex in t because it is the composition of the affine transformation $t\mapsto t\lambda_i$ with the convex function $-\log(x)$, where $x=1+t\lambda_i$. Affine transformations preserve this convexity and so each of these terms in the sum is convex $\implies$ the sum of these convex functions is convex so $-\sum_{i=1}^{n}\log\left(1+t\lambda_i\left(\mathbf{X}_0^{-\frac{1}{2}}\mathbf{Z}\mathbf{X}_0^{-\frac{1}{2}}\right)\right)$ is convex in t (i.e. along this line.) Thus, this proves that $\mathbf{X}\mapsto\log\det(\mathbf{X}^{-1})$ is convex as we have shown it is convex for any line (defined by a starting point and direction.)

2 Introduction to Convex Duality

We now begin our study of Convex Duality. The term duality comes from "dual space" notated by $X^*=$ all linear functions on a vector space $X=\mathbb{R}^n$ where $X^*\cong X$. Stated again, this means that each linear function is isomorphic to a vector representation in X.
The most fundamental duality result is known as *Separation Theorem*. We present it below.

Theorem 2.1 (Separation Theorem)**.** *Let $C\subset X$ be a closed, convex set and let $x_0\notin C$. Then, there exists a function $x^*\in X^*$ and a constant $\delta>0$ s.t. $\forall x\in C$:*

$$\langle x^*,x_0\rangle+\delta\leq\langle x^*,x\rangle.$$

Intuitively, Separation Theorem is telling us that for any point outside of our convex set C, we can "separate" it via some linear function $x^*\in X^*$ be at least δ-distance away from any point in the convex set. Separation Theorem is a foundational result of functional analysis and has led to many important results, including Farkas' Lemma which we will state and prove:

Theorem 2.2 (Farkas' Lemma)**.** *Let $A\in\mathbb{R}^{m\times n}, b\in\mathbb{R}^m$. Exactly one of the following is true:*

1. *There exists $x\geq 0$ such that $Ax=b$.*
2. *There exists y such that $A^Ty\geq 0$ and $b^Ty<0$.*

First note that this statement is saying that exactly one of the following is true. Before proving this statement, let us take a second to appreciate this lemma: Farkas' Lemma allows us to dramatically reduce the amount of work we need to expend in verifying that a given solution to a set of linear equations (i.e. a solution to u^* to $Au=b$) exists. We proceed with the proof.
Pf: Let $\mathcal{A}=$ conic hull of the columns of A, where the columns are given by $\{a_1,\ldots,a_n\}$. Recall that $\mathcal{A}=\{Ax : x\geq 0\}$. We prove this statement by casework:

1. Case One: $\exists\ x\geq 0$ s.t. $Ax=b$.

   This is our first case, and also the first statement in Farkas' Lemma. Note however that $\exists\ x\geq 0$ s.t. $Ax=b\iff b\in\mathcal{A}$. Thus, in our consideration of the other case, we are given $b\notin\mathcal{A}$.

2. Case Two: $b \notin \mathcal{A}$

   We apply Separation Theorem with conic hull $\mathcal{A}$ as our closed, convex set and $x_0 = b \notin \mathcal{A} \implies \exists\ x^* = y$ s.t. $\forall a \in \mathcal{A}, \langle y, b\rangle < \langle y, a\rangle$. Furthermore, note that $0 \in \mathcal{A} \implies$ (by Separation Theorem) $b^T y < \langle y, 0\rangle = 0$ as $b \notin \mathcal{A}$. This proves $b^T y < 0$. Finally because b is separated from $\mathcal{A}$ by hyperplane defined by y we have that $\forall x_i \geq 0, b^T y < 0 \leq \Sigma x_i \langle a_i, y\rangle \iff \forall i, a_i^T y \geq 0$. Because $a_i^T y = A^T y \geq 0$, we have entirely proved the second case in Farkas Lemmas'.

This completes our proof of Farkas' Lemma using Separation Theorem. We present one last (motivating) example.

Example 2.3. Take C, D as some disjoint closed bounded convex sets. This means there exists $x^* \in X^*$ and $\delta > 0$ such that

$$\langle x^*, c\rangle + \delta \leq \langle x^*, d\rangle \quad \forall c \in C \text{ and } \forall d \in D.$$

This is the main idea behind Support Vector Machines (SVMs), in which we want to find some separating hyperplane of two convex hulls corresponding to different high-dimensional training data of two different classes.

Pf: Let us define set $C - D = \{c - d : c \in C, d \in D\}$. It is a nice exercise to prove that $C - D$ is a closed, convex set. Furthermore, note $0 \notin C - D$ and so applying Separation Theorem with our choice $x_0 = 0 \notin C - D$, we have that $\exists\ x^*$ and $\exists\ \delta > 0$ s.t. $\forall y \in C - D, \langle x^*, y\rangle + \delta \leq \langle x^*, 0\rangle$. This shows that for any arbitrary choice of $c \in C$ and $d \in D$, $\langle x^*, c - d\rangle + \delta \leq \langle x^*, 0\rangle \implies \langle x^*, c\rangle + \delta \leq \langle x^*, d\rangle$.