EE 381V: Large Scale Optimization **Fall 2012**

Lecture 3 — September 06

Lecturer: Caramanis & Sanghavi *Scribe: Samer Chucri & Somsubhra Barik*

3.1 Topics covered

- Projection onto a Convex Set
- Separation of Convex Sets
- Unconstrained Optimization : Gradient Descent

In the last lecture, we began our discussion of projection onto a convex set. We review and then complete that discussion in the next section. After that, we use this to develop the notion of separation – a fundamental concept in convex analysis and convex optimization – and then use this to obtain what is known as an "outer description" of closed convex sets.

Finally, in the last section of this lecture, we began our foray into algorithms for convex optimization, and begin our discussion of Gradient Descent. However, it makes more sense for these notes to postpone this discussion until Lecture 4, in order to keep the topics together and more coherent.

3.2 Projection onto Convex Sets

Recall from the last lecture, that we defined the projection of a vector $\in \mathbb{R}^n$ onto a hyperplane $\mathcal{H}$ to be that vector given by

$$v^* \quad = \quad \arg\min_{v \in \mathcal{H}} \|x - v\|_2.$$

Note that we could in principle define this minimization with respect to any distance function. But the Euclidean norm, $\|\cdot\|_2$, used here, results in a "projection" which justifies its name, namely, the solution is characterized by an orthogonality condition: The vector v^* is the solution iff

$$\langle x - v^*, v^* - v \rangle \quad = \quad 0 \qquad \forall v \in \mathcal{H}.$$

Recall that we proved this in the last lecture, using the Pythagorean Theorem.

Next, we define the projection onto a given convex set $\mathcal{X} \subseteq \mathbb{R}^n$ analogously, using the same optimization problem:

$$\mathrm{Proj}_{\mathcal{X}}(x) \quad \stackrel{\triangle}{=} \quad \arg\min_{v \in \mathcal{X}} \|x - v\|. \tag{3.1}$$

Recall that in the last lecture, we showed that this is well-defined, namely, that as long as $\mathcal{X}$ is convex, then $\text{Proj}_{\mathcal{X}}(x)$ is unique. Note that this is not the case in general. For instance, if $\mathcal{X}$ is the sphere $\{x : \|x\|_2 = 1\}$, then every point in $\mathcal{X}$ is equidistant to the origin. We recap the statement of this result here. After showing uniqueness, we prove a variational characterization of the optimal solution, similar to the orthogonality condition for projection onto an affine manifold given above.

We prove uniqueness by proving something more general: the minimization of a strictly convex function has a unique minimizer. We repeat the definition of strong convexity given in the last lecture, as well as the statement of the uniqueness result.

Definition 1: *A function $f : \mathbb{R}^n \to \mathbb{R}$ is called* strictly convex *if and only if for any $x_1, x_2 \in \mathbb{R}^n$ such that $x_1 \neq x_2$ and $\lambda \in (0, 1)$,*

$$f(\lambda x_1 + (1 - \lambda x_2)) \quad < \quad \lambda f(x_1) + (1 - \lambda) f(x_2).$$

Strict convexity can be also characterized by the first order gradient condition:

$$f(x_2) \quad > \quad f(x_1) + \langle \nabla f(x_1), x_2 - x_1 \rangle.$$

Geometrically, this means that the straight line segment joining any two points on the function lies strictly above the function except for its endpoints. See, e.g., Fig. 3.1. For twice differentiable functions, you can check that a function is strictly convex as long as the Hessian, $\nabla^2 f$, is positive definite (strictly positive eigenvalues). Thus, in particular, the squared Euclidean distance function, $f(x) = \|x - y\|_2^2$, is strictly convex.

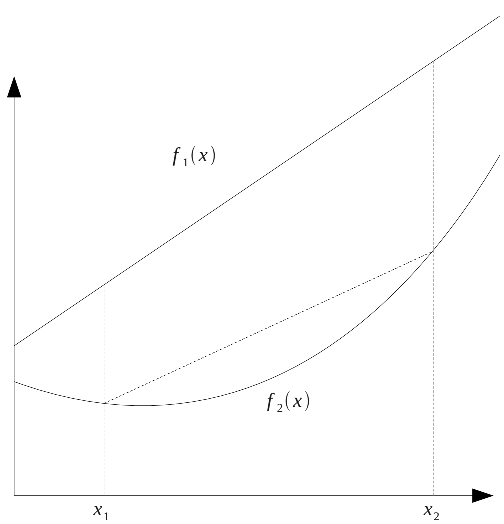

Figure 3.1. This graph shows two functions: $f_1(x)$ is convex but not strictly so, where as $f_2(x)$ is a strictly convex function.

Proposition 1. *If $\mathcal{X} \subseteq \mathbb{R}^n$ is a convex set and the function $f : \mathbb{R}^n \to \mathbb{R}$ is strictly convex, then the problem*

$$\begin{array}{rl} \min & f(x) \\ s.t. & x \in \mathcal{X} \end{array}$$

has a unique solution if it has any solutions (proved in the previous lecture).

Proposition 2. *Let $\mathcal{X} \in \mathbb{R}^n$ be a* closed *and* convex *set and let $x \notin \mathcal{X}$. Then a solution to the problem in (3.1) exists and is unique.*

Remark 1. *As mentioned in the last lecture, $\mathcal{X}$ is required to be closed in the proposition to ensure that the solution to the minimization problem (3.1) is contained in $\mathcal{X}$. Otherwise, if $\mathcal{X}$ is open (e.g., $\mathcal{X} = \{x : \|x\|_2 < 1\}$), then the projection of $x \notin \mathcal{X}$ onto $\mathcal{X}$ will be a point on the boundary of $\mathcal{X}$ defined by $\mathcal{X}_B = \{x : \|x\|_2 = 1\}$ where $\mathcal{X}_B \not\subseteq \mathcal{X}$. Also, in order for the solution to exist, we need an additional condition. Requiring that the intersection of $\mathcal{X}$ with a sub-level set of f to be compact, suffices. For an example where existence of an optimal solution can fail, consider the convex function $f(x) = 1/x$, and the closed convex set $\mathcal{X} = \mathbb{R}_+$.*

Now we can prove the variational characterization of projection.

Proposition 3. For any $x \notin \mathcal{X}$, $v^* = \mathrm{Proj}_{\mathcal{X}}(x)$ if and only if $\langle x - v^*, v - v^* \rangle \leq 0, \ \forall v \in \mathcal{X}$.

The statement is illustrated in Fig. 3.2.

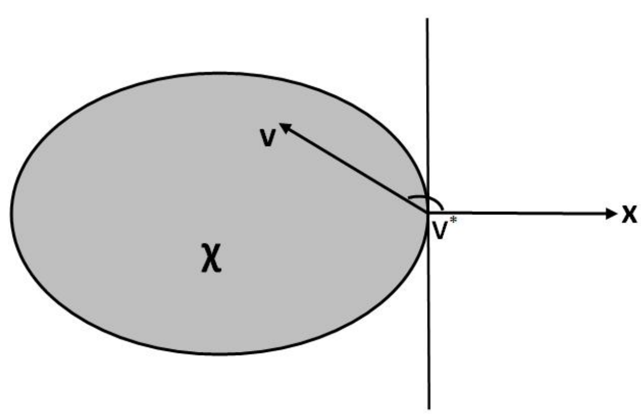

Figure 3.2. Projection onto a Convex Set

Proof: ($\Rightarrow$) Let $v^* = \mathrm{Proj}_{\mathcal{X}}(x)$ for a given $x \notin \mathcal{X}$, that is, suppose that v^* is the unique solution to the optimization problem. Let $v \in \mathcal{X}$ be such that $v \neq v^*$. Let $\alpha \in (0,1)$. Since $\mathcal{X}$ is convex, $(1-\alpha)v^* + \alpha v = v^* + \alpha(v - v^*) \in \mathcal{X}$. By the (assumed) optimality of v^*, we must have:

$$\begin{aligned} \|x - v^*\|^2 &\leq \|x - (v^* + \alpha(v - v^*))\|^2 \\ &= \|x - v^*\|^2 + \alpha^2\|v - v^*\|^2 - 2\alpha\langle x - v^*, v - v^*\rangle \\ \Rightarrow \quad \langle x - v^*, v - v^*\rangle &\leq \frac{\alpha}{2}\|v - v^*\|^2. \end{aligned} \tag{3.2}$$

Now, note that (3.2) holds for $\forall\alpha \in (0,1)$. Since the RHS of (**??**) can be made arbitrarily small for a given v, the LHS cannot be strictly positive. Thus we conclude, as desired:

$$\langle x - v^*, v - v^*\rangle \;\leq\; 0, \;\forall v \in \mathcal{X}.$$

($\Leftarrow$) Let $v^* \in \mathcal{X}$ be such that $\langle x - v^*, v - v^*\rangle \leq 0, \;\forall v \in \mathcal{X}$. We show that it must be the optimal solution. Let $v \in \mathcal{X}$ and $v \neq v^*$.

$$\begin{aligned} \|x - v\|^2 - \|x - v^*\|^2 &= \|x - v^* + (v^* - v)\|^2 - \|x - v^*\|^2 \\ &= \cancel{\|x - v^*\|^2} + \|v - v^*\|^2 - 2\langle x - v^*, v - v^*\rangle - \cancel{\|x - v^*\|^2} \\ &> 0. \end{aligned}$$

Hence, v^* is the optimal solution to the optimization problem, and thus $v^* = \mathrm{Proj}_{\mathcal{X}}(x)$ by definition. $\square$

3.3 Separation of Convex Sets

We use the projection result to obtain an outer representation of convex sets.

We show that closed convex sets can have dual representations as intersection of half spaces as shown in Figure 3.3. Note that non-convex sets (closed or not) do not have a similar representation. It is straightforward to see this by a picture, e.g., see Fig. 3.4. To prove such an impossibility result, recall that the intersection of any (finite or not) collection of convex sets is again convex. See also Fig. 3.4

The first step to showing this, is demonstrating that if $\mathcal{X}$ is a closed convex set, and $x \notin \mathcal{X}$, then there is a simple "proof" or "certificate" or this fact: there is a hyperplane so that $\mathcal{X}$ lies on one side, and x on the other. Recall that for a hyperplane $\mathcal{H} = \{x : \langle s, x\rangle = b\}$, positive and negative halfspaces are defined as $\mathcal{H}^+ = \{x : \langle s, x\rangle \geq b\}$ and $\mathcal{H}^- = \{x : \langle s, x\rangle \leq b\}$. The interior of $\mathcal{H}^+$ is defined as $\mathrm{Int}(\mathcal{H}^+) = \{x : \langle s, x\rangle > b\}$. $\mathrm{Int}(\mathcal{H}^-)$ is defined similarly.

Proposition 4. Let $\mathcal{X}$ be a closed convex set and $x \notin \mathcal{X}$. Then, there exists a hyperplane $\mathcal{H}$ such that $\mathcal{X} \subseteq \mathcal{H}^-$ and $x \in \mathrm{int}(\mathcal{H}^+)$.

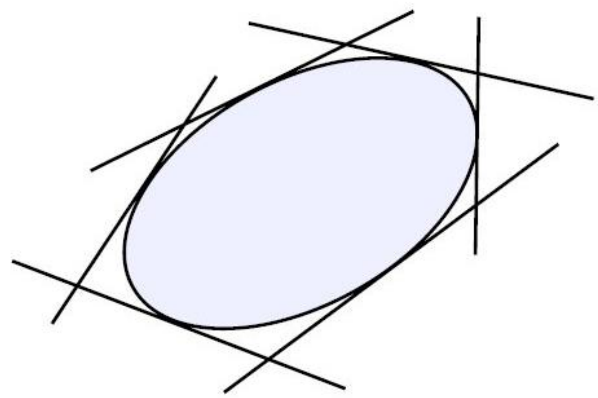

Figure 3.3. Alternative representation of Convex Sets: Intersection of half spaces

Proof: Let $v^* = \text{Proj}_{\mathcal{X}}(x)$. By Proposition 3, we have that since v^* is the projection, it must satisfy:

$$\begin{aligned} \langle x - v^*, v - v^* \rangle &\leq 0, \quad \forall v \in \mathcal{X} \\ \Rightarrow \quad \langle x - v^*, v \rangle &\leq \langle x - v^*, v^* \rangle, \end{aligned}$$

i.e., every $v \in \mathcal{X}$ satisfies $\langle s, v \rangle \leq b$, where, $s = x - v^*$ is the normal vector and $b = \langle x - v^*, v^* \rangle$ is the intercept of some hyperplane $\mathcal{H} = \{x : \langle s, x \rangle = b\}$. Hence, $\mathcal{X} \subseteq \mathcal{H}^-$.
Again, $v^* \in \mathcal{X}$ because $\mathcal{X}$ is closed. Since $x \notin \mathcal{X}$,

$$\begin{aligned} \|x - v^*\|^2 = \langle x - v^*, x - v^* \rangle &> 0 \\ \Rightarrow \quad \langle x - v^*, x \rangle &> \langle x - v^*, v^* \rangle \end{aligned}$$

i.e., x satisfies $\langle s, x \rangle > b$ with s and b as defined above. Hence, $x \in \text{int}(\mathcal{H}^+)$. □

We can now use this to show that if $\mathcal{X}$ is a closed convex set, it is the intersection of all half-spaces that contain it. We give one relevant definition:

Definition 1. *Given a set $\mathcal{X}$, a hyperplane $\mathcal{H}$ such that $\mathcal{X} \subseteq \mathcal{H}^+$ or $\mathcal{X} \subseteq \mathcal{H}^-$ is called a* supporting hyperplane.

Corollary 3.1. *Let $\mathcal{X} \subseteq \mathbb{R}^n$ be a closed convex set. Then, $\mathcal{X} = \bigcap_{\mathcal{H}^- \supseteq \mathcal{X}} \mathcal{H}^-$, i.e., $\mathcal{X}$ is the intersection of all such halfspaces containing $\mathcal{X}$.*

Proof: The "$\subseteq$" inclusion is immediate: $\mathcal{X}$ is a subset of all such half spaces $\mathcal{H}^-$, hence $\mathcal{X} \subseteq \bigcap_{\mathcal{H}^- \supseteq \mathcal{X}} \mathcal{H}^-$.

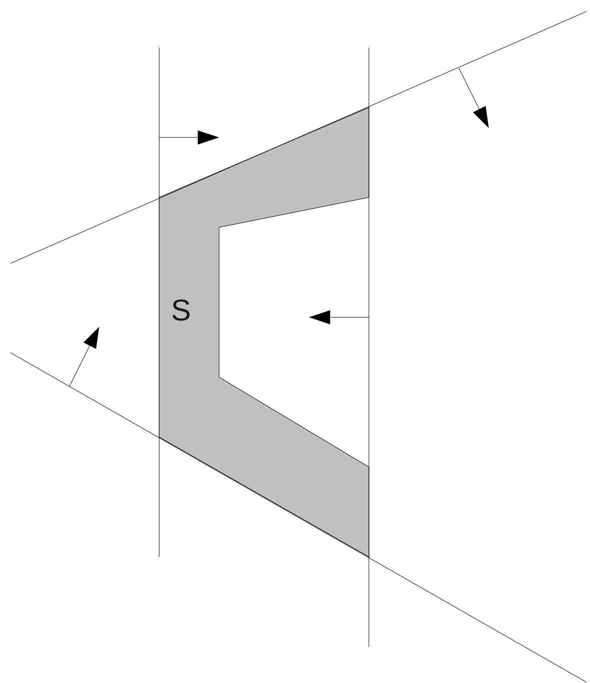

Figure 3.4. A non-convex set S cannot be described as an intersection of half-spaces. Only $\mathrm{cl}(\mathrm{Conv}(S))$, the closed convex hull, can be recovered from the outer description.

For the reverse inclusion, consider any point $x \notin \mathcal{X}$. We show that this x is excluded from $\bigcap_{\mathcal{H}^- \supseteq \mathcal{X}} \mathcal{H}^-$ (thus we are effectively proving that the complement of the LHS is contained in the complement of the RHS). Since $\mathcal{X}$ is closed and convex, and $x \notin \mathcal{X}$, we can apply our previous result to conclude that there is a hyperplane $\mathcal{H}_x$ such that $\mathcal{X} \subseteq \mathcal{H}_x^-$, while $x \in \mathrm{int}\mathcal{H}_x^+$. But then by definition, the hyperplane $\mathcal{H}_x$ is included in the intersection making up the right hand side, and therefore $x \notin \bigcap_{\mathcal{H}^- \supseteq \mathcal{X}} \mathcal{H}^-$, which is precisely what we needed to show. $\square$

We complete this portion of the lecture by defining several notions of separation. Let $\mathcal{X}_1$ and $\mathcal{X}_2$ be nonempty convex subsets of $\mathbb{R}^n$.

Definition 2. *The sets $\mathcal{X}_1$ and $\mathcal{X}_2$ are said to be* properly separated *if there exists a hyperplane $\mathcal{H}$ such that,*

$$\begin{aligned} \mathcal{X}_1 &\subseteq \mathcal{H}^-, \\ \mathcal{X}_2 &\subseteq \mathcal{H}^+ \qquad \textit{and} \\ \mathcal{X}_1 \cup \mathcal{X}_2 &\subseteq \mathcal{H} \end{aligned}$$

Definition 3. *The sets $\mathcal{X}_1$ and $\mathcal{X}_2$ are said to be* strictly separated *if there exists a hyperplane $\mathcal{H}$ such that*

$$\begin{aligned} \mathcal{X}_1 &\subseteq \operatorname{Int}(\mathcal{H}^-) \quad \textit{and} \\ \mathcal{X}_2 &\subseteq \operatorname{Int}(\mathcal{H}^+) \end{aligned}$$

Definition 4. *The sets $\mathcal{X}_1$ and $\mathcal{X}_2$ are said to be* strongly separated *if there exist a hyperplane $\mathcal{H}$ and $\epsilon > 0$ such that*

$$\begin{aligned} \mathcal{X}_1 + \mathcal{B}_\epsilon(0) &\subseteq \operatorname{Int}(\mathcal{H}^-) \quad \textit{and} \\ \mathcal{X}_2 + \mathcal{B}_\epsilon(0) &\subseteq \operatorname{Int}(\mathcal{H}^+) \end{aligned} \tag{3.3}$$

where, $\mathcal{B}_\epsilon(0) = \{z \in \mathbb{R}^n : \|z\| < \epsilon\}$. In other words, an open ball of radius ϵ around any $x \in \mathcal{X}_1$ will be contained in interior of $\mathcal{H}^-$ and similarly for $\mathcal{X}_2$. See Fig. (3.5))

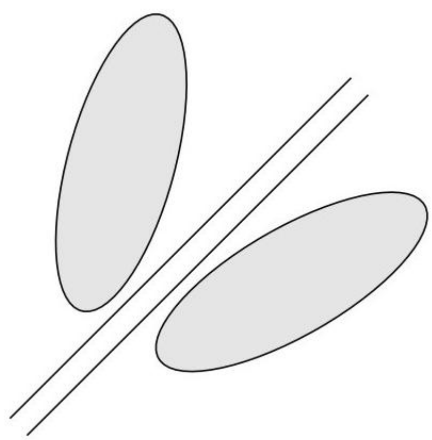

Figure 3.5. Strong Separation of convex sets

Example: Let $\mathcal{X}_1 = \{(x_1, x_2) \in \mathbb{R}^2 : x_1 < 0, x_2 > 1/|x_1|\}$ and $\mathcal{X}_2 = \{(x_1, x_2) \in \mathbb{R}^2 : x_1 > 0, x_2 > 1/x_1\}$. Then the hyperplane $\mathcal{H} = \{x_1 = 0\}$ strictly separates $\mathcal{X}_1$ and $\mathcal{X}_2$ but not strongly (see Figure (3.6))

3.4 Unconstrained optimization via descent method

At this point, we transitioned into discussion algorithms, and in particular, descent methods. The scribing for this part of the class is postponed until Lecture 4, where it will be more coherent and cohesive.

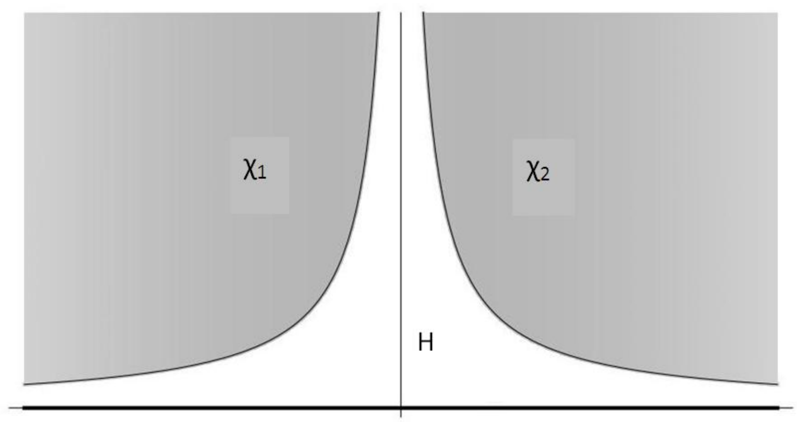

Figure 3.6. $\mathcal{X}_1$ and $\mathcal{X}_2$ are strictly but not strongly separated by $\mathcal{H}$.