ISEN 629: Engineering Optimization

Lecture 10

Sergiy Butenko

Industrial and Systems Engineering
Texas A&M University

Fall 2007

Convex sets

- Next we will look at constrained minimization problem in the form
$$\min_{x \in Q} f(x),$$
where $Q \subset \mathbb{R}^n$.
- Given the vectors $x_1, \ldots, x_m$ in Euclidean space $\mathbb{R}^n$ and real numbers $\lambda_i \geq 0$ with $\sum_{i=1}^{m} \lambda_i = 1$, the vector sum $\lambda_1 x_1 + \ldots + \lambda_m x_m$ is called a **convex combination** of these points.
- For example, the convex combination of two points is the line segment between these two points, and the convex combination of three non–colinear points is a triangle.

Convex sets

- A subset C of $\mathbb{R}^n$ is said to be **convex** if for every x_1, $x_2 \in C$ and $\lambda \in \mathbb{R}$, $0 \leq \lambda \leq 1$, we have $\lambda x_1 + (1-\lambda)x_2 \in C$.
- The geometric interpretation: for any two points of C, the line segment joining them lies entirely in C.

Theorem

A subset of $\mathbb{R}^n$ is convex iff it contains all the convex combinations of its elements.

Convex sets

Theorem (2.2.4)

Let $Q_1 \subseteq \mathbb{R}^n$ and $Q_2 \subseteq \mathbb{R}^m$ be convex sets and $\mathcal{A}(x)$ be a linear operator: $\mathcal{A}(x) = Ax + b : \mathbb{R}^n \to \mathbb{R}^m$. Then all sets below are convex:

1. *Intersection* $(m = n) : Q_1 \cap Q_2 = \{x \in \mathbb{R}^n : x \in Q_1, x \in Q_2\}$.
2. *Sum* $(m = n) : Q_1 + Q_2 = \{z = x + y : x \in Q_1, y \in Q_2\}$.
3. *Direct sum (Cartesian product):* $Q_1 \times Q_2 = \{(x, y) \in \mathbb{R}^{n+m} : x \in Q_1, y \in Q_2\}$.
4. *Conic hull:* $\mathcal{K}(Q_1) = \{z \in \mathbb{R}^n : z = \beta x, x \in Q_1, \beta \geq 0\}$.
5. *Convex hull:*
$$Conv(Q_1, Q_2) = \{z \in \mathbb{R}^n : \; z = \alpha x + (1-\alpha)y, \; x \in Q_1, y \in Q_2, \alpha \in [0,1]\}.$$
6. *Affine image:* $\mathcal{A}(Q_1) = \{y \in \mathbb{R}^m : y = \mathcal{A}(x), x \in Q_1\}$.
7. *Inverse affine image:* $\mathcal{A}^{-1}(Q_2) = \{x \in \mathbb{R}^n : \mathcal{A}(x) \in Q_2\}$.

Convex sets

Examples:

- A **hyperplane** H in $\mathbb{R}^n$ is a set of the form
$$H = \{x \in \mathbb{R}^n : c^T x = b\},$$
where $c \in \mathbb{R}^n \setminus \{0\}$ and $b \in \mathbb{R}$. Similarly we define the closed halfspaces
$$H_+ = \{x \in \mathbb{R}^n : c^T x \geq b\},$$
$$H_- = \{x \in \mathbb{R}^n : c^T x \leq b\}.$$
- It is easy to see that H, H_+, H_- are all convex sets since linear function is convex.

Convex sets

Examples:

- The intersection of finitely many halfspaces is called a **polyhedron**.
- Using matrix notation we can define a polyhedron to be the set of points $P = \{x \in \mathbb{R}^n : Ax \leq b\}$, where A is an $m \times n$ matrix and $b \in \mathbb{R}^m$.
- A bounded polyhedron is called a **polytope**.
- Polyhedron (polytope) is convex as an intersection of convex sets.
- Ellipsoid $\{x \in \mathbb{R}^n : x^T Ax \leq r^2\}$, where $A = A^T \succeq 0$, is convex since the function $x^T Ax$ is convex.

Convex sets

- A nonempty subset V of $\mathbb{R}^n$ is called a *(linear) subspace* if the following conditions hold:
  (i) if $x, y \in V$, then $x + y \in V$;
  (ii) if $x \in V$ and $r \in \mathbb{R}$, then $rx \in V$.
- We say that a set of vectors $S = \{v_1, \dots, v_m\} \subset V$ spans V if every vector $v \in V$ is a linear combination of vectors from S: $v = \sum_{i=1}^{m} c_i v_i$, where each $c_i \in \mathbb{R}$.

Convex sets

- The vectors in S are are linearly independent if $\sum_{i=1}^{m} c_i v_i = 0 \Rightarrow c_i = 0, i = 1, \dots, m$.
- If S spans V and its elements $v_1, \dots, v_m$ are linearly independent we call S a basis of V.
- The dimension of the subspace V, $\dim(V)$ is the number of vectors in some basis S of V.
- Alternatively, a linear subspace V can be defined as solution set of the homogeneous system of linear equations $Cx = 0$, where C is a $m \times n$ matrix. Here the dimension of V is $n - rank(C)$.

Convex sets

- An affine subspace A of $\mathbb{R}^n$ is a linear subspace V translated by some vector u, i.e., $A = \{x \in \mathbb{R}^n : x = u + v, v \in V$. We have dim($A$)=dim($V$).
- An affine set is a set that contains the line between any two distinct points in the set.
- Equivalently we can define A as the solution set of the nonhomogeneous linear system $Cx = b$.
- Note that a hyperplane in $\mathbb{R}^n$ is an affine subspace of dimension $n - 1$.

Convex sets and functions

Lemma (2.2.5)
If $f(x)$ is a convex function then for any $\beta \in \mathbb{R}$ its lower level set

$$\mathcal{L}_f(\beta) = \{x \in \mathbb{R}^n : f(x) \leq \beta\}$$

is either convex or empty.
Note that the convexity of $\mathcal{L}_f(\beta)$ does not imply that f is convex.

Lemma (2.2.6)
$f(x)$ is a convex function if and only if its epigraph

$$\mathcal{E}_f = \{(x, \tau) \in \mathbb{R}^{n+1} : f(x) \leq \tau\}$$

is a convex set.

Constrained convex problems

Theorem (2.2.5)
Let $f \in \mathcal{F}^1(\mathbb{R}^n)$ and Q be a closed convex set. A point x^* is a solution of the problem $\min_{x \in Q} f(x)$ if and only if

$$f'(x^*)^T(x - x^*) \geq 0$$

for all $x \in Q$.

Constrained convex problems

Theorem (2.2.6)
Let $f \in \mathcal{S}^1_\mu(\mathbb{R}^n)$ and Q be a closed convex set. Then there exists a unique solution x^* of problem $\min_{x \in Q} f(x)$.

Proof: For $x_0 \in Q$, consider the set $\bar{Q} = \{x \in Q : f(x) \leq f(x_0)\}$.

$$f(x_0) \geq f(x) \geq f(x_0) + f'(x_0)^T(x - x_0) + \frac{\mu}{2}\|x - x_0\|^2,$$

so $\|x - x_0\|^2 \leq -\frac{2}{\mu} f'(x_0)^T(x - x_0) \leq \frac{2}{\mu}\|f'(x_0)\|\|x - x_0\|$, and thus $\bar{Q}$ is bounded. Note that our problem is equivalent to the problem $\min_{x \in \bar{Q}} f(x)$, therefore it has a solution x^*. Assume there is another solution $\bar{x}$. Then

$$\begin{aligned} f^* &= f(\bar{x}) \geq f(x^*) + f'(x^*)^T(\bar{x} - x^*) + \tfrac{\mu}{2}\|\bar{x} - x^*\|^2 \\ &\geq f^* + \|\bar{x} - x^*\|^2 \quad \Rightarrow \quad \bar{x} = x^*. \qquad \square \end{aligned}$$

Gradient mapping

Definition

For a fixed $\gamma > 0$, denote by

$$\begin{aligned} x_Q(\bar{x}, \gamma) &= \arg\min_{x \in Q} [f(\bar{x}) + f'(\bar{x})^T (x - \bar{x}) + \tfrac{\gamma}{2} \|x - \bar{x}\|^2], \\ g_Q(\bar{x}, \gamma) &= \gamma(\bar{x} - x_Q(\bar{x}, \gamma)). \end{aligned}$$

We call $g_Q(\bar{x}, \gamma)$ the gradient mapping of f on Q.

Note that for $Q = \mathbb{R}^n$ we have

$$x_Q(\bar{x}, \gamma) = \bar{x} - \frac{1}{\gamma} f'(\bar{x}), \; g_Q(\bar{x}, \gamma) = f'(\bar{x}),$$

so $1/\gamma$ can be viewed as the step size of the step from $\bar{x}$ to $x_Q(\bar{x}, \gamma)$.

Gradient mapping

Theorem (2.2.7)

Let $f \in \mathcal{S}^{1,1}_{\mu,L}(\mathbb{R}^n), \gamma \geq L$ and $\bar{x} \in \mathbb{R}^n$. Then for any $x \in Q$ we have

$$f(x) \geq f(x_Q(\bar{x}, \gamma)) + g_Q(\bar{x}, \gamma)^T (x - \bar{x}) + \frac{1}{2\gamma} \|g_Q(\bar{x}, \gamma)\|^2 + \frac{\mu}{2} \|x - \bar{x}\|^2.$$

- $x = \bar{x}$:
$$f(x_Q(\bar{x}, \gamma)) \leq f(\bar{x}) - \frac{1}{2\gamma} \|g_Q(\bar{x}, \gamma)\|^2.$$
- $x = x^*$ (and thus, $f(x_Q(\bar{x}, \gamma)) \geq f(x^*)$):
$$g_Q(\bar{x}, \gamma)^T (\bar{x} - x^*) \geq \frac{1}{2\gamma} \|g_Q(\bar{x}, \gamma)\|^2 + \frac{\mu}{2} \|x^* - \bar{x}\|^2.$$

(compare to the "unconstrained" properties of the gradient:

$$f(x - \frac{1}{L} f'(x)) \leq f(x) - \frac{1}{2L} \|f'(x)\|^2; \;\; f'(x)^T (x - x^*) \geq \frac{1}{L} \|f'(x)\|^2).$$

Minimization over simple sets

- We consider the problem
$$\min_{x \in Q} f(x),$$
where $f \in \mathcal{S}^{1,1}_{\mu,L}(\mathbb{R}^n)$ and Q is a closed convex set.
- We assume that Q is simple enough for computing the gradient mapping explicitly.
- Then the gradient method scheme for the constrained problem is:
  0. Choose $x_0 \in Q$.
  1. k-th iteration ($k \geq 0$):
$$x_{k+1} = x_k - h g_Q(x_k, L).$$

Minimization over simple sets

Theorem (2.2.8)

Let $f \in \mathcal{S}^{1,1}_{\mu,L}(\mathbb{R}^n)$. If in the above scheme $h = \frac{1}{L}$, then

$$\|x_k - x^*\|^2 \leq \left(1 - \frac{\mu}{L}\right)^k \|x_0 - x^*\|^2.$$

Proof: Denote by

$$r_k = \|x_k - x^*\|, \; g_Q = g_Q(x_k, L).$$

Using the inequality

$$g_Q(\bar{x}, \gamma)^T (\bar{x} - x^*) \geq \frac{1}{2\gamma} \|g_Q(\bar{x}, \gamma)\|^2 + \frac{\mu}{2} \|x^* - \bar{x}\|^2,$$

we obtain

$$\begin{aligned} r_{k+1}^2 &= \|x_k - x^* - h g_Q\|^2 = r_k^2 - 2h g_Q^T (x_k - x^*) + h^2 \|g_Q\|^2 \\ &\leq (1 - h\mu) r_k^2 + h \left(h - \tfrac{1}{L}\right) \|g_Q\|^2 \\ &= \left(1 - \tfrac{\mu}{L}\right) r_k^2. \end{aligned}$$

□

Optimal schemes

- Optimal schemes are derived similarly to the unconstrained case.
- Again, we define an estimate sequence, etc.
- The same convergence properties as in the unconstrained case.

Optimal schemes

0. Choose $x_0 \in \mathbb{R}^n$ and $\alpha_0 \in (0, 1)$.
   Set $y_0 = x_0$ and $q = \mu/L$.
1. k-th iteration ($k \geq 0$):
   a) Compute $f(y_k)$ and $f'(y_k)$. Set
   $$x_{k+1} = x_Q(y_k, L).$$
   b) Compute $\alpha_{k+1} \in (0, 1)$ from equation
   $$\alpha_{k+1}^2 = (1 - \alpha_{k+1})\alpha_k^2 + q\alpha_{k+1}.$$
   Set $\beta_k = \frac{\alpha_k(1-\alpha_k)}{\alpha_k^2 + \alpha_{k+1}}$,
   $$y_{k+1} = x_{k+1} + \beta_k(x_{k+1} - x_k).$$