C&O 355
Lecture 15

[N. Harvey](#)

Topics

- Subgradient Inequality
- Characterizations of Convex Functions
- Convex Minimization over a Polyhedron
- (Mini)-KKT Theorem
- Smallest Enclosing Ball Problem

Subgradient Inequality

- **Prop:** Suppose f : $\mathbb{R} \to \mathbb{R}$ is differentiable.
  Then f is convex iff
  $$f(y) \geq f(x) + f'(x)(y-x) \qquad \forall x,y \in \mathbb{R}$$
- **Proof:**
  $\Leftarrow$: See Notes Section 3.2.
  $\Rightarrow$: Exercise for Assignment 4. $\square$

Convexity and Second Derivative

- **Prop:** Suppose f : $\mathbb{R} \to \mathbb{R}$ is twice-differentiable. Then f is convex iff f''(x)$\geq$0 $\forall$x$\in\mathbb{R}$.
- **Proof:** See Notes Section 3.2.

Subgradient Inequality in $\mathbb{R}^n$

- **Prop:** Suppose $f : \mathbb{R}^n \to \mathbb{R}$ is differentiable.
  Then f is convex iff
  $$f(y) \geq f(x) + \nabla f(x)^T(y-x) \qquad \forall x,y \in \mathbb{R}$$
- **Proof:**
  $\Leftarrow$: Exercise for Assignment 4.
  $\Rightarrow$: See Notes Section 3.2. $\square$

Minimizing over a Convex Set

- **Prop:** Let $C \subseteq \mathbb{R}^n$ be a convex set.
  Let $f : \mathbb{R}^n \rightarrow \mathbb{R}$ be convex and differentiable.
  Then x minimizes f over C iff $\nabla f(x)^T(z-x) \geq 0 \;\forall z \in C$.
- **Proof:** $\Leftarrow$ direction

  Direct from subgradient inequality.

$$f(z) \underbrace{\geq}_{\text{Subgradient inequality}} f(x) + \nabla f(x)^T(z-x) \underbrace{\geq}_{\text{Our hypothesis}} f(x)$$

Minimizing over a Convex Set

- **Prop:** Let $C \subseteq \mathbb{R}^n$ be a convex set.
  Let $f : \mathbb{R}^n \to \mathbb{R}$ be convex and differentiable.
  Then x minimizes f over C iff $\nabla f(x)^T (z-x) \geq 0 \ \forall z \in C$.

- **Proof:** $\Rightarrow$ direction

  Let x be a minimizer, let $z \in C$ and let $y = z-x$.

  Recall that $\nabla f(x)^T y = f'(x;y) = \lim_{t \to 0} \frac{f(x+ty)-f(x)}{t}$.

  If limit is negative then we have $f(x+ty)<f(x)$ for some $t \in [0,1]$, contradicting that x is a minimizer.

  So the limit is non-negative, and $\nabla f(x)^T y \geq 0$. ■

Positive Semidefinite Matrices (again)

- Assume M is symmetric
- **Old definition:** M is PSD if $\exists$V s.t. M = $V^T V$.
- **New definition:** M is PSD if $y^T M y \geq 0 \; \forall y \in \mathbb{R}^n$.
- **Claim:** Old $\Rightarrow$ New.
- **Proof:** $y^T M y = y^T V^T V y = \|Vy\|^2 \geq 0$.
- **Claim:** New $\Rightarrow$ Old.
- **Proof:** Based on spectral decomposition of M.

Convexity and Hessian

- **Prop:** Let $f:\mathbb{R}^n \rightarrow \mathbb{R}$ be a C^2-function.
  Let $H(x)$ denote the Hessian of f at point x.
  Then f is convex iff $H(x)$ is PSD $\forall x \in \mathbb{R}^n$.

- **Proof:** $\Rightarrow$ direction

  Fix $y \in \mathbb{R}^n$. Consider function $g_y(\alpha) = f(x+\alpha y)$.

  Convexity of g_y follows from convexity of f.

  Thus $g_y''(0) \geq 0$. (convexity & 2nd derivative)

  Fact: $g_y''(0) = y^T H(x)\, y$ (stated in Lecture 14)

  So $y^T H(x)\, y \geq 0\ \forall y \in \mathbb{R}^n \ \Rightarrow\ H(x)$ is PSD.

Convexity and Hessian

- **Prop:** Let $f:\mathbb{R}^n \to \mathbb{R}$ be a C^2-function.
  Let $H(x)$ denote the Hessian of f at point x.
  Then f is convex iff $H(x)$ is PSD $\forall x \in \mathbb{R}^n$.
- **Proof:** $\Leftarrow$ direction. Fix $x,y \in \mathbb{R}^n$.
  Define $g : [0,1] \to \mathbb{R}$ by $g(\alpha) = f(x+\alpha(y-x))$.
  For all α, $g''(\alpha) = (y-x)^T H(x+\alpha(y-x))(y-x) \geq 0$.
  (The equality was stated in Lecture 14. The inequality holds since H is PSD.)
  So g is convex. (By Prop "Convexity and Second Derivative")
  $f((1-\alpha)x+\alpha y) = g(\alpha) \leq (1-\alpha)g(0) + \alpha g(1)$
  $= (1-\alpha)f(x) + \alpha f(y).$ ■

Hessian Example

- **Example:**

  Let M be a symmetric n×n matrix. Let $z \in \mathbb{R}^n$.

  Define $f : \mathbb{R}^n \to \mathbb{R}$ by $\boxed{f(x) = x^T M x - x^T z.}$

  Note $f(x) = \Sigma_i \Sigma_j M_{i,j} x_i x_j - \Sigma_i x_i z_i$.

  Taking partial derivatives $\nabla f(x)_i = 2 \Sigma_j M_{i,j} x_j - z_i$.

  So $\boxed{\nabla f(x) = 2Mx - z.}$

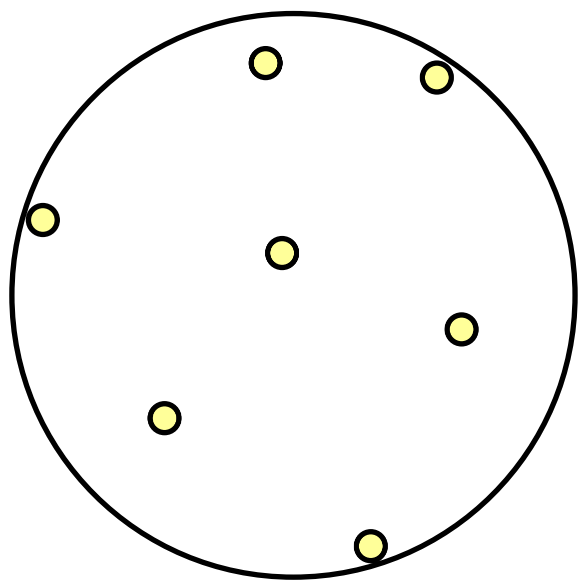
  Recall the Hessian at x is $H(x) = \nabla(\nabla f(x))$.

  So $\boxed{H(x) = 2M.}$

Smallest Ball Problem

- Let $\{p_1,\ldots,p_n\}$ be points in $\mathbb{R}^d$.
  Find (unique!) ball of smallest volume (not an ellipsoid!) that contains all the p_i's.
- In other words, we want to solve:

$$\min \{ r : \exists y \in \mathbb{R}^d \text{ s.t. } p_i \in B(y,r) \ \forall i \}$$



Smallest Ball Problem

- Let $\{p_1,\ldots,p_n\}$ be points in $\mathbb{R}^d$.
  Find (unique!) ball of smallest volume (not an ellipsoid!) that contains all the p_i's.
- In other words, we want to solve:
  $$\min \{ r : \exists y \in \mathbb{R}^d \text{ s.t. } p_i \in B(y,r)\ \forall i \}$$
- We will formulate this as a convex program.
- In fact, our convex program will be of the form
  $$\min \{ f(x) : Ax=b,\ x \geq 0 \},\ \text{ where } f \text{ is convex.}$$
  Minimizing a convex function over an (equality form) polyhedron
- To solve this, we will need optimality conditions for convex programs.

(Mini)-KKT Theorem

Theorem: Let $f:\mathbb{R}^n \to \mathbb{R}$ be a convex, C^2 function.
Let $x \in \mathbb{R}^n$ be a feasible solution to the convex program

$$\min \{ f(x) : Ax=b,\ x \geq 0 \}$$

Then x is optimal iff $\exists y \in \mathbb{R}^m$ s.t.

1) $\nabla f(x)^T + y^T A \geq 0$,
2) For all j, if $x_j > 0$ then $\nabla f(x)_j + y^T A_j = 0$.

j^{th} column of A

- Proven by Karush in 1939 (his Master's thesis!), and by Kuhn and Tucker in 1951.

(Mini)-KKT Theorem

Theorem: Let $f:\mathbb{R}^n \to \mathbb{R}$ be a convex, C^2 function.

Let $x \in \mathbb{R}^n$ be a feasible solution to the convex program

$$\min \{ f(x) : Ax=b,\ x \geq 0 \}$$

Then x is optimal iff $\exists y \in \mathbb{R}^m$ s.t.

1) $\nabla f(x)^T + y^T A \geq 0$,
2) For all j, if $x_j > 0$ then $\nabla f(x)_j + y^T A_j = 0$.

Special Case: (Strong LP Duality)

Let $f(x) = -c^T x$. (So the convex program is $\max \{ c^T x : Ax=b,\ x \geq 0 \}$)

Then x is optimal iff $\exists y \in \mathbb{R}^m$ s.t.

1) $-c^T + y^T A \geq 0$,
2) For all j, if $x_j > 0$ then $-c_j + y^T A_j = 0$.

y is feasible for Dual LP

Complementary Slackness holds

Theorem: Let $f:\mathbb{R}^n \to \mathbb{R}$ be a convex, C^2 function.

Let $x \in \mathbb{R}^n$ be a feasible solution to the convex program

$$\min \{ f(x) : Ax=b,\ x \geq 0 \}$$

Then x is optimal iff $\exists y \in \mathbb{R}^m$ s.t.

1) $\nabla f(x)^T + y^T A \geq 0$,
2) For all j, if $x_j > 0$ then $\nabla f(x)_j + y^T A_j = 0$.

Proof: $\Leftarrow$ direction. Suppose such a y exists. Then

$$(\nabla f(x)^T + y^T A)\, x = 0.$$ (Just like complementary slackness)

For any feasible $z \in \mathbb{R}^n$, we have

$$(\nabla f(x)^T + y^T A)\, z \geq 0.$$

Subtracting these, and using $Ax=Az=b$, we get

$$\nabla f(x)^T (z-x) \geq 0 \quad \forall \text{ feasible } z.$$

So x is optimal. (By earlier proposition “Minimizing over a Convex Set”)

Theorem: Let $f:\mathbb{R}^n \to \mathbb{R}$ be a convex, C^2 function.

Let $x \in \mathbb{R}^n$ be a feasible solution to the convex program

$$\min \{ f(x) : Ax=b, x \geq 0 \}$$

Then x is optimal iff $\exists y \in \mathbb{R}^m$ s.t.

1) $\nabla f(x)^T + y^T A \geq 0$,
2) For all j, if $x_j > 0$ then $\nabla f(x)_j + y^T A_j = 0$.

Proof: $\Rightarrow$ direction. Suppose x is optimal. Let $c = -\nabla f(x)$.

Then $\nabla f(x)^T (z-x) \geq 0 \;\; \Rightarrow \;\; c^T z \leq c^T x$ for all feasible points z.

By our earlier proposition "Minimizing over a Convex Set"

Theorem: Let $f:\mathbb{R}^n \to \mathbb{R}$ be a convex, C^2 function.
Let $x \in \mathbb{R}^n$ be a feasible solution to the convex program

$$\min \{ f(x) : Ax=b,\ x \geq 0 \}$$

Then x is optimal iff $\exists y \in \mathbb{R}^m$ s.t.

1) $\nabla f(x)^T + y^T A \geq 0$,
2) For all j, if $x_j>0$ then $\nabla f(x)_j + y^T A_j = 0$.

Proof: $\Rightarrow$ direction. Suppose x is optimal. Let $c=-\nabla f(x)$.
Then $\nabla f(x)^T(z-x) \geq 0 \;\Rightarrow\; c^T z \leq c^T x$ for all feasible points z.
So x is optimal for the LP $\max \{ c^T x : Ax=b,\ x \geq 0 \}$.
So there is an optimal solution y to dual LP $\min \{ b^T y : A^T y \geq c \}$.
So $\nabla f(x)^T + y^T A = -c^T + y^T A \geq 0 \quad \Rightarrow \quad$ (1) holds.

Furthermore, x and y are both optimal so C.S. holds.

$\Rightarrow$ whenever $x_j>0$, the j^{th} dual constraint is tight
$\Rightarrow$ $y^T A_j = c_j \quad \Rightarrow \quad$ (2) holds. ∎

Smallest Ball Problem

- Let $P=\{p_1,\ldots,p_n\}$ be points in $\mathbb{R}^d$.
  Let Q be $d \times n$ matrix s.t. $Q_i = p_i$.
  Let $z \in \mathbb{R}^n$ satisfy $z_i = p_i^T p_i$.
  Define $f : \mathbb{R}^n \to \mathbb{R}$ by $f(x) = x^T Q^T Q x - x^T z$.
- **Claim 1:** f is convex.
- Consider the convex program

$$\min \{ f(x) : \textstyle\sum_j x_j = 1,\ x \geq 0 \}.$$

- **Claim 2:** This program has an optimal solution x.
- **Claim 3:** Let $p^* = Qx$ and $r = \sqrt{-f(x)}$. Then $P \subset B(p^*, r)$.
- **Claim 4:** $B(p^*, r)$ is the smallest ball containing P.

- Let $P=\{p_1,\ldots,p_n\}$ be points in $\mathbb{R}^d$.
  Let Q be $d \times n$ matrix s.t. $Q_i=p_i$.
  Let $z \in \mathbb{R}^n$ satisfy $z_i=p_i^T p_i$.
  Define $f : \mathbb{R}^n \rightarrow \mathbb{R}$ by $f(x) = x^T Q^T Q x - x^T z$.
- **Claim 1:** f is convex.
- **Proof:**
  By our earlier example, the Hessian of f at x is
  $$H(x) = 2 \cdot Q^T Q.$$
  This is positive semi-definite.

  By our proposition "Convexity and Hessian",
  f is convex. $\square$

- Let $P=\{p_1,\ldots,p_n\}$ be points in $\mathbb{R}^d$.
  Let Q be $d\times n$ matrix s.t. $Q_i=p_i$.
  Let $z\in\mathbb{R}^n$ satisfy $z_i=p_i^T p_i$.
  Define $f : \mathbb{R}^n \to \mathbb{R}$ by $f(x) = x^T Q^T Q x - x^T z$.
- Consider the convex program

$$\min \{ f(x) : \textstyle\sum_j x_j = 1,\ x\geq 0 \}.$$

- **Claim 2:** This program has an optimal solution.
- **Proof:**
  The objective function is continuous.

  The feasible region is a bounded polyhedron, and hence compact.

  By Weierstrass' Theorem, an optimal solution exists. $\square$

- Let Q be d×n matrix s.t. $Q_i = p_i$.
  Let $z \in \mathbb{R}^n$ satisfy $z_i = p_i^T p_i$.
  Define $f : \mathbb{R}^n \to \mathbb{R}$ by $f(x) = x^T Q^T Q x - x^T z$.
- Let x be an optimal solution of $\min \{ f(x) : \sum_j x_j = 1, x \geq 0 \}$
  Let $p^* = Qx$ and $r^2 = -f(x) = \sum_j x_j p_j^T p_j - p^{*T} p^*$.
- **Claim 3:** The ball $B(p^*, r)$ contains P.
- **Proof:** Note that $\nabla f(x)^T = 2x^T Q^T Q - z^T$.
  By KKT, $\exists y \in \mathbb{R}$ s.t. $2p_j^T p^* - p_j^T p_j + y \geq 0 \ \forall j$.
  Furthermore, equality holds $\forall j$ s.t. $x_j > 0$.
  So $y = \sum_j x_j y = \sum_j x_j p_j^T p_j - 2\sum_j x_j p_j^T p^* = \sum_j x_j p_j^T p_j - 2p^{*T} p^*$.

  So $y + p^{*T} p^* = \sum_j x_j p_j^T p_j - p^{*T} p^* = -f(x) \Rightarrow r = \sqrt{y + p^{*T} p^*}$

- Let $p^* = Qx$ and $r^2 = -f(x) = \sum_j x_j p_j^T p_j - p^{*T}p^*$.
- **Claim 3:** The ball $B(p^*,r)$ contains P.
- **Proof:** Note that $\nabla f(x) = 2x^TQ^TQ - z$.

  By KKT, $\exists y \in \mathbb{R}$ s.t. $2p_j^Tp^* - p_j^Tp_j + y \geq 0\ \forall j$.

  Furthermore, equality holds $\forall j$ s.t. $x_j > 0$.

  $\Rightarrow\ r = \sqrt{y + p^{*T}p^*}$

  It remains to show that $B(p^*,r)$ contains P.

  This holds iff $\|p_j - p^*\| \leq r\ \forall j$.

  Now $\|p_j-p^*\|^2 = (p_j-p^*)^T(p_j-p^*)$

  $$= p^{*T}p^* - 2p_j^Tp^* + p_j^Tp_j$$

  $$\leq y + p^{*T}p^* = r^2\ \forall j. \qquad \square$$

- **Claim 4:** $B(p^*, r)$ is the smallest ball containing P.
- **Proof:** See Textbook.