Proof Script for SS18 Convex Optimization Lecture*

Last updated: May 6, 2018

1 Convex Analysis

Theorem 1.1 (separation of convex sets)**.** *Let C_1, C_2 be nonempty convex subsets in $\mathbb{E}$ such that $C_1 \cap C_2 = \emptyset$ and C_1 is open. Then there exists a hyperplane separating C_1 and C_2, i.e. $\exists v \in \mathbb{E},\ v \neq 0,\ \alpha \in \mathbb{R}$ such that*

$$\left\langle v, u^1 \right\rangle \geq \alpha \geq \left\langle v, u^2 \right\rangle, \quad \forall u^1 \in C_1,\ u^2 \in C_2.$$

Proof. (i) Claim: Let $C \subset \mathbb{E}$ be closed, convex set, and $w \in \mathbb{E} \backslash C$. Then $\exists v \in \mathbb{E},\ v \neq 0,\ \alpha \in \mathbb{R}$ s.t. $\langle v, w \rangle > \alpha \geq \langle v, u \rangle\ \forall u \in C$.

Consider the projection of w onto C, i.e. set $u^* := \arg\min_{u \in C} \frac{1}{2}\|u - w\|^2$ or, equivalently, let $\langle u - u^*, u^* - w \rangle \geq 0\ \forall u \in C$.

Now set $v := w - u^* \neq 0$. Then $\forall u \in C$, we have $\langle v, w \rangle = \langle w - u^*, w \rangle = \|w - u^*\|^2 + \langle w - u^*, u^* \rangle \geq \|w - u^*\|^2 + \langle w - u^*, u \rangle = \|v\|^2 + \langle v, u \rangle$. Set $\alpha := \sup\{\langle v, u \rangle : u \in C\}$. Note $\alpha < \infty$ since $\langle v, u \rangle \leq \langle v, u^* \rangle\ \forall u \in C$. Thus $\langle v, w \rangle > \alpha \geq \langle v, u \rangle\ \forall u \in C$, which proves the claim.

(ii) Let C_1 be an open, convex subset of $\mathbb{E}$, and $C_2 = \{\bar{w}\}$ with $\bar{w} \in \mathbb{E} \backslash C_1$. Since $\mathbb{E} \backslash C_1$ is closed, $\exists w^k \in \mathbb{E} \backslash \operatorname{cl} C_1$ s.t. $w^k \to \bar{w}$. For each w^k, by (i), $\exists v^k \in \mathbb{E}$ with $\|v^k\| \equiv 1$ s.t. $\left\langle v^k, w^k \right\rangle \leq \left\langle v^k, u^1 \right\rangle\ \forall u^1 \in C_1 \subset \operatorname{cl} C_1$. Hence $v^k \to \bar{v} \in \mathbb{E}$ along a subsequence s.t. $\|\bar{v}\| = 1$ and $\langle \bar{v}, \bar{w} \rangle \leq \left\langle \bar{v}, u^1 \right\rangle\ \forall u^1 \in C_1$.

(iii) Consider C_2 as a general convex subset of $\mathbb{E}$. Set $C := C_2 - C_1 = \{u^2 - u^1 : u^1 \in C_1,\ u^2 \in C_2\}$. Note that C is a convex, open set, and $0 \notin C$. By (ii), $\exists \bar{v} \in \mathbb{E}$ with $\|\bar{v}\| = 1$ s.t. $\left\langle -\bar{v}, u^2 - u^1 \right\rangle \geq \langle -\bar{v}, 0 \rangle = 0$ or, equivalently, $\left\langle \bar{v}, u^1 \right\rangle \geq \left\langle \bar{v}, u^2 \right\rangle\ \forall u^1 \in C_1,\ u^2 \in C_2$. Set $\alpha := \sup\{\left\langle \bar{v}, u^2 \right\rangle : u^2 \in C_2\}$, then we conclude that $\left\langle \bar{v}, u^1 \right\rangle \geq \alpha \geq \left\langle \bar{v}, u^2 \right\rangle\ \forall u^1 \in C_1,\ u^2 \in C_2$. □

Theorem 1.2. *A proper convex function $J : \mathbb{E} \to \overline{\mathbb{R}}$ is locally Lipschitz at any $u \in \operatorname{rint} \operatorname{dom} J$.*

Proof. Throughout the proof, we consider $J : \operatorname{aff} \operatorname{dom} J \to \overline{\mathbb{R}}$.

(i) Claim: If $M = \sup\{J(v) : v \in B_\epsilon(u)\} < \infty$ with $\epsilon > 0$, then J is locally Lipschitz at u.

First, by convexity of J we have $\forall v \in B_\epsilon(u) : J(v) \geq 2J(u) - J(2u - v) \geq 2J(u) - M$. Thus, $\sup\{|J(v)| : v \in B_\epsilon(u)\} \leq M + 2|J(u)|$.

Next, we show J is Lipschitz on $B_{\epsilon/2}(u)$. Let $v, w \in B_{\epsilon/2}(u)$ be given. Take $z \in B_\epsilon(u)$ s.t. $w = (1 - t)v + tz$ for some $t \in [0, 1]$ and $\|z - v\| \geq \epsilon/2$. By convexity, $J(w) - J(v) \leq t(J(z) - J(v)) \leq 2t(M - J(u))$. Since $t(z - v) = w - v$, we have $t = \|w - v\|/\|z - v\| \leq 2\|w - v\|/\epsilon$ and $J(w) - J(v) \leq (4(M - J(u))/\epsilon)\|w - v\|$. Analogously, one can show $J(v) - J(w) \leq (4(M - J(u))/\epsilon)\|w - v\|$. Hence, J is Lipschitz on $B_{\epsilon/2}(u)$ with modulus $4(M - J(u))/\epsilon$.

*Please report typos to: tao.wu@tum.de

(ii) Let $u \in \operatorname{rint} \operatorname{dom} J$ and $n = \dim(\operatorname{aff} \operatorname{dom} J)$. Then by Carathéodory's theorem, $\exists\{\alpha^i\}_{i=1}^{n+1} \subset (0,1)$, $\{u^i\}_{i=1}^{n+1} \subset \operatorname{dom} J$ s.t. $u = \sum_{i=1}^{n+1} \alpha^i u^i$, $\sum_{i=1}^{n+1} \alpha^i = 1$, i.e. u belongs to the interior of the convex hull of $\{u^i\}_{i=1}^{n+1}$. Thus one can apply (i) to assert that J is locally Lipschitz at u. □

Theorem 1.3. *For any proper convex function $J : \mathbb{E} \to \overline{\mathbb{R}}$, if $u^* \in \operatorname{dom} J$ is a local minimizer of J, then it is also a global minimizer.*

Proof. By the definition of a local minimizer, $\exists \epsilon > 0$ s.t. $J(u^*) \leq J(u)\ \forall u \in B_\epsilon(u^*)$. For the sake of contradiction, assume $\exists \bar{u} \in \mathbb{E}$ s.t. $J(\bar{u}) < J(u^*)$. By convexity of J, we have $J(\alpha\bar{u} + (1-\alpha)u^*) \leq J(u^*) - \alpha(J(u^*) - J(\bar{u})) < J(u^*)\ \forall \alpha \in (0,1]$. This violates the local optimality of u^* as $\alpha \to 0^+$. □

Theorem 1.4. *Any proper function $J : \mathbb{E} \to \overline{\mathbb{R}}$, which is bounded from below, coercive, and lsc, has a (global) minimizer.*

Proof. Let $\{u^k\}$ be an infimizing sequence for J, i.e. $\lim_{k\to\infty} J(u^k) = \inf_{u\in\mathbb{E}} J(u) > -\infty$. Since $\{J(u^k)\}$ is uniformly bounded from above, by coercivity of J, $\{u^k\}$ is uniformly bounded. By compactness, $u^k \to u^*$ along a subsequence. Since J is lsc, we have $J(u^*) \leq \liminf_{k\to\infty} J(u^k) = \inf_{u\in\mathbb{E}} J(u)$, which implies $J(u^*) = \inf_{u\in\mathbb{E}} J(u)$ or u^* is a minimizer of J. □

Theorem 1.5. *The minimizer of a strictly convex function $J : \mathbb{E} \to \overline{\mathbb{R}}$ is unique.*

Proof. Let $u, v \in \mathbb{E}$ be two (global) minimizers s.t. $u \neq v$ and $J(u) = J(v) = J^*$. By strict convexity of J, $J(\alpha u + (1-\alpha)v) < \alpha J(u) + (1-\alpha)J(v) = J^*$ for all $\alpha \in (0,1)$, which contradicts the global optimality of u and v. □

Theorem 1.6. *Let $J : \mathbb{E} \to \overline{\mathbb{R}}$ be a convex function. Then ∂J is a monotone operator, i.e. $\forall u^1, u^2 \in \operatorname{dom} J,\ p^1 \in \partial J(u^1),\ p^2 \in \partial J(u^2)$:*

$$\left\langle p^1 - p^2, u^1 - u^2 \right\rangle \geq 0.$$

Proof. By applying the definition of subdifferential at arbitrarily given $u^1, u^2 \in \operatorname{dom} J$, we have

$$\begin{aligned} J(u^2) &\geq J(u^1) + \left\langle p^1, u^2 - u^1 \right\rangle, \\ J(u^1) &\geq J(u^2) + \left\langle p^2, u^1 - u^2 \right\rangle. \end{aligned}$$

Adding the two inequalities yields $\left\langle p^1 - p^2, u^1 - u^2 \right\rangle \geq 0$. □

Theorem 1.7. *Let $J : \mathbb{E} \to \overline{\mathbb{R}}$ be a convex function. Then for any $u \in \operatorname{int} \operatorname{dom} J$, $\partial J(u)$ is a nonempty, compact, and convex subset.*

Proof. (i) nonemptiness. Since $(u, J(u)) \notin \operatorname{int} \operatorname{epi} J$, by Theorem 1.1 we have $\exists (p, -\alpha) \in \mathbb{E} \times \mathbb{R}$ s.t. $(p, -\alpha) \neq (0,0)$, $\alpha \geq 0$ by our choice, and $\langle (p, -\alpha), (u - v, J(u) - J(v)) \rangle \geq 0\ \forall v \in \operatorname{dom} J$. In fact, we must have $\alpha > 0$ since otherwise $p = 0$. Thus, we conclude that $p/\alpha \in \partial J(u)$.

(ii) boundedness. By Theorem 1.2, J is locally Lipschitz at u with modulus L_u. Let $p \in \partial J(u)$ be fixed. For any $h \in (\operatorname{dom} J) - u$ whenever $\|h\|$ is sufficiently small, we have $\langle p, h \rangle \leq J(u + h) - J(u) \leq L_u \|h\|$. This holds true only if $\|p\| \leq L_u$, which implies boundedness of $\partial J(u)$.

(iii) closedness. Let $v \in \mathbb{E}$ be arbitrarily fixed and $p^k \to p^*$ where each $p^k \in \partial J(u)$. Then $\forall k : J(v) - J(u) \geq \left\langle p^k, v - u \right\rangle$. By continuity, $J(v) - J(u) \geq \langle p^*, v - u \rangle$ when passing $k \to \infty$. Since v can be arbitrary, we assert $p^* \in \partial J(u)$.

(iv) convexity. Let $v \in \mathbb{E}$ be arbitrarily fixed, and $p, q \in \partial J(u)$. Then we have

$$J(v) \geq J(u) + \langle p, v - u \rangle ,$$
$$J(v) \geq J(u) + \langle q, v - u \rangle .$$

Hence, $\forall 0 \leq \alpha \leq 1 : J(v) \geq J(u) + \langle \alpha p + (1-\alpha) q, v - u \rangle$, i.e. $\alpha p + (1-\alpha) q \in \partial J(u)$. □

Theorem 1.8. *Let $J : \mathbb{E} \to \overline{\mathbb{R}}$ be a proper, convex, lsc function. Then ∂J is a closed set-valued map, i.e. $p^* \in \partial J(u^*)$ whenever*

$$\exists (u^k, p^k) \to (u^*, p^*) \in (\operatorname{dom} J) \times \mathbb{E} \;\text{ s.t. } p^k \in \partial J(u^k) \;\forall k.$$

Proof. Let $v \in \mathbb{E}$ be arbitrarily fixed. For each k, $p^k \in \partial J(u^k) \;\Rightarrow\; J(v) \geq J(u^k) + \langle p^k, v - u^k \rangle$. Passing $k \to \infty$, we have $\langle p^k, v - u^k \rangle \to \langle p^*, v - u^* \rangle$ and $J(u^*) \leq \liminf_{k\to\infty} J(u^k)$. Hence, $J(u^*) + \langle p^*, v - u^* \rangle \leq \liminf_{k\to\infty} \{J(u^k) + \langle p^k, v - u^k \rangle\} \leq J(v)$. Since v can be arbitrary, $p^* \in \partial J(u^*)$. □

Theorem 1.9. *Given any proper convex function $J : \mathbb{E} \to \overline{\mathbb{R}}$, the sufficient and necessary condition for u^* being a (global) minimizer for J is: $0 \in \partial J(u^*)$.*

Proof. (i) sufficiency. $0 \in \partial J(u^*) \;\Rightarrow\; J(u) \geq J(u^*) + \langle 0, u - u^* \rangle = J(u^*) \;\forall u \in \mathbb{E}$.
(ii) necessity. $J(u^*) \leq J(u) \;\forall u \in \mathbb{E} \;\Rightarrow\; J(u^*) + \langle 0, u - u^* \rangle \leq J(u) \;\forall u \;\Rightarrow\; 0 \in \partial J(u^*)$. □

Theorem 1.10 (Fenchel-Young inequality)**.** *For any $J : \mathbb{E} \to \overline{\mathbb{R}}$ and $(u, p) \in \mathbb{E} \times \mathbb{E}$, we have*

$$J(u) + J^*(p) \geq \langle u, p \rangle .$$

The equality holds iff $p \in \partial J(u)$ for $(u, p) \in \operatorname{dom} J \times \operatorname{dom} J^*$.

Proof. (i) $J(u) + J^*(p) \geq \langle u, p \rangle$ follows directly from the definition of convex conjugate. (ii) The equality holds only if $(u, p) \in \operatorname{dom} J \times \operatorname{dom} J^*$. Moreover, $p \in \partial J(u)$ is the sufficient and necessary condition for $\min_{u \in \mathbb{E}} \{J(u) - \langle u, p \rangle\}$. □

Theorem 1.11 (order reversing)**.** *For any $J_1, J_2 : \mathbb{E} \to \overline{\mathbb{R}}$, we have $J_1^* \leq J_2^*$ whenever $J_1 \geq J_2$.*

Proof. Given any (u, p), we have $\langle u, p \rangle - J_1(u) \leq \langle u, p \rangle - J_2(u)$. Taking supremum over u on both sides yields $J_1^*(p) \leq J_2^*(p)$. □

Theorem 1.12. *Let $J : \mathbb{E} \to \overline{\mathbb{R}}$, and $J^{**} = (J^*)^*$ be the biconjugate of J. In general:*

1. *$J^{**}(\cdot) \leq J(\cdot)$.*
2. *J^* is convex and lsc.*

If J is proper, convex, and lsc, then:

3. *$J^{**}(\cdot) = J(\cdot)$.*
4. *$p \in \partial J(u)$ iff $u \in \partial J^*(p)$.*

Proof. (1) Since $J^{**}(u) = \sup_p\{\langle p, u\rangle - J^*(p)\}$ and, by Theorem 1.10, $\langle p, u\rangle - J^*(p) \le J(u)\ \forall p$, we assert $J^{**}(\cdot) \le J(\cdot)$.

(2) (i) convexity. Let $p, q \in \mathbb{E}$, $0 \le \alpha \le 1$. Then $J^*(\alpha p+(1-\alpha)q) = \sup_u\{\langle u, \alpha p + (1-\alpha)q\rangle - J(u)\} \le \sup_u\{\langle \alpha u, p\rangle - \alpha J(u)\} + \sup_u\{\langle (1-\alpha)u, q\rangle - (1-\alpha)J(u)\} = \alpha J^*(p) + (1-\alpha)J^*(q)$.

(ii) lsc. Note $\operatorname{epi} J^* = \{(p,\alpha) \in \mathbb{E}\times\mathbb{R} : \langle u, p\rangle - J(u) \le \alpha\ \forall u\} = \cap_u \operatorname{epi}\Phi_u$ where $\Phi_u(\cdot) = \langle u, \cdot\rangle - J(u)$. Since each $\operatorname{epi}\Phi_u$ and any arbitrary intersection of closed sets is closed, $\operatorname{epi} J^*$ is closed and hence J^* is lsc.

(3) For the sake of contradiction, assume $\exists \bar u \in \operatorname{dom} J^{**}$ s.t. $J(\bar u) > J^{**}(\bar u)$. Let $d \in (0, J(\bar u) - J^{**}(\bar u))$ be fixed. Since $(\bar u, J(\bar u) - d) \notin \operatorname{epi} J$ and $\operatorname{epi} J$ is convex and closed, by Theorem 1.1, $\exists(\bar p, -1) \in \mathbb{E}\times\mathbb{R}$ s.t. $\langle(\bar p,-1),(\bar u, J(\bar u)-d)\rangle \ge \langle(\bar p,-1),(u,\alpha)\rangle\ \forall(u,\alpha) \in \operatorname{epi} J$. In particular, $\langle \bar p, \bar u\rangle - J(\bar u) + d \ge \langle \bar p, u\rangle - J(u)\ \forall u \in \operatorname{dom} J$. Hence, $\langle \bar p, \bar u\rangle - J(\bar u) + d \ge J^*(\bar p) \ge \langle \bar p, \bar u\rangle - J^{**}(\bar u)$ by Theorem 1.10. Thus we have $J^{**}(\bar u) + d \ge J(\bar u)$ as a contradiction to our assumption.

(4) $p \in \partial J(u) \iff J(u) + J^*(p) = \langle u, p\rangle \iff J^{**}(u) + J^*(p) = \langle u, p\rangle \iff u \in \partial J^*(p)$. □

Theorem 1.13. *Assume that $J : \mathbb{E} \to \overline{\mathbb{R}}$ is proper, convex, and lsc. Then J is μ-strongly convex iff J^* is $\frac{1}{\mu}$-Lipschitz differentiable.*

Proof. (only if) Let $p \in \partial J(u)$ be arbitrarily given. By μ-strong convexity of J, we have

$$J(v) \ge J(u) + \langle p, v-u\rangle + \frac{\mu}{2}\|v-u\|^2 \quad \forall v. \tag{1}$$

Then $\forall q : J^*(q) = \sup_v\{\langle q, v\rangle - J(v)\} \le \sup_v\{\langle q, v\rangle - J(u) - \langle p, v-u\rangle - \frac{\mu}{2}\|v-u\|^2\} = \langle q, u\rangle - J(u) + \sup_v\{\langle q-p, v-u\rangle - \frac{\mu}{2}\|v-u\|^2\} = \langle q, u\rangle - J(u) + \frac{1}{2\mu}\|q-p\|^2 = \langle p, u\rangle - J(u) + \langle q-p, u\rangle + \frac{1}{2\mu}\|q-p\|^2 = J^*(p) + \langle q-p, u\rangle + \frac{1}{2\mu}\|q-p\|^2$. Here we have used the identity $\langle p, u\rangle - J(u) = J^*(p)$. We have actually derived $\lim_{q\to p} \|J^*(q) - J^*(p) - \langle q-p, u\rangle\|/\|q-p\| = 0$, which asserts that J^* is (Frechét-)differentiable at p with $\nabla J^*(p) = u$.

Finally we show ∇J^* is $\frac{1}{\mu}$-Lipschitz. Let $u = \nabla J^*(p)$, $v = \nabla J^*(q)$, or equivalently $p \in \partial J(u)$, $q \in \partial J(v)$. Then by (1) we have

$$\begin{aligned} J(v) &\ge J(u) + \langle p, v-u\rangle + \frac{\mu}{2}\|v-u\|^2,\\ J(u) &\ge J(v) + \langle q, u-v\rangle + \frac{\mu}{2}\|u-v\|^2. \end{aligned}$$

Adding the above two inequalities, we obtain $\mu\|u-v\|^2 \le \langle p-q, u-v\rangle \le \|p-q\|\|u-v\|$ and thus $\|u-v\| \le \frac{1}{\mu}\|p-q\|$.

(if) Note that $J^*(q) = J^*(p) + \int_0^1 \langle \nabla J^*(p+s(q-p)), q-p\rangle\, ds = J^*(p) + \langle \nabla J^*(p), q-p\rangle + \int_0^1 \langle \nabla J^*(p+s(q-p)) - \nabla J^*(p), q-p\rangle\, ds \le J^*(p) + \langle \nabla J^*(p), q-p\rangle + \frac{1}{2\mu}\|q-p\|^2$. Let $p \in \partial J(u) \iff u = \nabla J^*(p)$. Then $J^*(q) \le J^*(p) + \langle q-p, u\rangle + \frac{1}{2\mu}\|q-p\|^2$. Taking the convex conjugate on both sides, we deduce $J(v) = J^{**}(v) \ge \sup_q\{\langle q, v\rangle - [J^*(p) + \langle q-p, u\rangle + \frac{1}{2\mu}\|q-p\|^2]\} = -J^*(p) + \langle p, v\rangle + \frac{\mu}{2}\|v-u\|^2 = J(u) + \langle p, v-u\rangle + \frac{\mu}{2}\|v-u\|^2$. □

Theorem 1.14 (weak duality)**.** *Let $K \in \mathbb{R}^{m\times n}$, and $F : \mathbb{R}^m \to \overline{\mathbb{R}}$, $G : \mathbb{R}^n \to \overline{\mathbb{R}}$ are proper, convex, and lsc. Then it holds that* $\inf_u\{F(Ku) + G(u)\} \ge \sup_p\{-G^*(-K^\top p) - F^*(p)\}$.

Proof. Let $\mathcal{L}(u,p) = \langle p, Ku\rangle - F^*(p) + G(u)$, then $\inf_u\{F(Ku) + G(u)\} = \inf_u \sup_p \mathcal{L}(u,p)$ and $\sup_p\{-G^*(-K^\top p) - F^*(p)\} = \sup_p \inf_u \mathcal{L}(u,p)$. It remains to verify $\inf_u \sup_p \mathcal{L}(u,p) \ge \sup_p \inf_u \mathcal{L}(u,p)$. For an arbitrarily fixed (u,p), we have $\sup_{p'} \mathcal{L}(u,p') \ge \mathcal{L}(u,p) \ge \inf_{u'} \mathcal{L}(u',p)$. Hence, the conclusion follows. □