THE FENCHEL–ROCKAFELLAR THEOREM

AIDAN BACKUS

1. Subdifferential calculus

Let X be a Banach space and let X^* be its dual space.

Definition 1.1. Let J be a convex function on X, $x_0 \in X$, and $x^* \in X^*$. Then x^* is a *subgradient* of J at x_0, if for every $x \in X$,

$$J(x) - J(x_0) \geq \langle x^*, x\rangle.$$

If J has a subgradient at x_0, then J is *subdifferentiable* at x_0.

Theorem 1.2. *Let J be a convex function on X and $x_0 \in X$. If J is continuous at x_0 and $-\infty < J(x_0) < +\infty$, then J is subdifferentiable at x_0.*

Proof. Let

$$\Gamma := \{(x, t) \in X \times \mathbf{R} : J(x) \leq t\}$$

be the epigraph of J. Since J is continuous at x_0 and convex, and $-\infty < J(x_0) < \infty$, Γ has nonempty and convex interior, Γ°, and $(x_0, J(x_0)) \notin \Gamma^\circ$. So by the Hanh-Banach theorem, there exists an affine hyperplane $H \subset X$ such that $(x_0, J(x_0)) \in H$ and $\Gamma^\circ \cap H = \emptyset$. Furthermore, there exists $(x^*, t^*) \in X^* \in \mathbf{R}$ such that (x^*, t^*) is an upwards-pointing conormal vector to H, in the sense that there exists $s \in \mathbf{R}$ such that

$$H = \{(x, t) \in X \times \mathbf{R} : \langle x^*, x\rangle + t^* t = s\}$$

and for any $(x, t) \in \Gamma$,

$$\langle x^*, x\rangle + t^* t \geq s.$$

Therefore, for any $x \in X$,

$$\langle x^*, x - x_0\rangle \geq -t^*(J(x) - J(x_0)).$$

So $-x^*/t^*$ is a subgradient of J at x_0. □

2. The Fenchel–Rockafellar theorem

Let $\Lambda : X \to Y$ be a continuous linear map of Banach spaces. The situation we have in mind is that there is a manifold M, with de Rham complex $\Omega^\bullet$, and we are considering the linear map

$$\mathrm{d} : W^{1,p}(M, \Omega^k) \to L^p(M, \Omega^{k+1}),$$

where $p \in [1, \infty)$, but there is no need to be so explicit at this stage.

Suppose that we are given a convex function J on $X \times Y$, and we want to minimize

$$F(x) := J(x, \Lambda x).$$

Let

$$J^*(x^*, y^*) = \sup_{(x,y)\in X\times Y} \langle x^*, x\rangle + \langle y^*, y\rangle - J(x, y)$$

be the convex conjugate of J, which is a convex function on the dual space, $X^* \times Y^*$.

Date: February 7, 2025.
2020 *Mathematics Subject Classification.* primary: 49Q20; secondary: 35J92, 49Q05, 37F34, 81T35.

Definition 2.1. The *dual problem* of minimizing F is the problem of maximizing

$$G(y^*) := -J^*(\Lambda^* y^*, -y^*).$$

A pair $(\underline{x}, \overline{y}^*) \in X \times Y^*$ satisfies the *duality relation* if

$$F(\underline{x}) = G(\overline{y}^*).$$

Lemma 2.2. *One has*

$$\sup_{y^* \in Y^*} G(y^*) \leq \inf_{x \in X} F(x).$$

Proof. Let $y^* \in Y^*$. Then

$$\begin{aligned} G(y^*) &= -J^*(\Lambda^* y^*, -y^*) \\ &= -\sup_{(x,y) \in X \times Y} [\langle \Lambda^* y^*, x\rangle - \langle y^*, y\rangle - J(x,y)] \\ &= \inf_{(x,y) \in X \times Y} [J(x,y) + \langle y^*, y\rangle - \langle \Lambda^* y^*, x\rangle] \\ &\leq \inf_{x \in X} [J(x, \Lambda x) + \langle y^*, \Lambda x\rangle - \langle \Lambda^* y^*, x\rangle] \\ &= \inf_{x \in X} F(x). \end{aligned}$$

□

I'm not aware that there's a standard name for the below technical condition, but it's important and annoying to write out every time.

Definition 2.3. (J, Λ) is *suitable* if F is not identically $-\infty$ and there exists $x_0 \in X$ such that $F(x_0) < +\infty$ and $y \mapsto J(x_0, y)$ is continuous at Λx_0.

Theorem 2.4 (Fenchel–Rockafellar theorem)**.** *Suppose that* (J, Λ) *is suitable. Then* G *has a maximum* $\overline{y}^*$ *such that*

$$G(\overline{y}^*) = \inf_{x \in X} F(x). \tag{2.1}$$

Proof. Let

$$h(y) := \inf_{x \in X} J(x, \Lambda x - y).$$

Since J is convex, h is convex. By suitability,

$$h(0) = \inf_{x \in X} F(x) \leq F(x_0) < +\infty.$$

Also by suitability, there exists $\varepsilon > 0$ and $M < +\infty$ such that if $|y| < \varepsilon$ then

$$h(y) \leq \inf_{x \in X} J(x, \Lambda x - y) \leq J(x_0, \Lambda x_0 - y) \leq M < +\infty.$$

Since h is convex, it follows that h is continuous at 0. So by Theorem 1.2, there exists $\overline{y}^* \in Y^*$ such that for every $y \in Y$,

$$h(y) \geq h(0) + \langle \overline{y}^*, y\rangle.$$

In other words,

$$\inf_{y \in Y} [h(y) - \langle \overline{y}^*, y\rangle] \geq h(0) = \inf_{x \in X} F(x).$$

Applying this inequality, we compute

$$\begin{aligned} G(\overline{y}^*) &= -J^*(\Lambda^*\overline{y}^*, -\overline{y}^*) \\ &= -\sup_{(x,y)\in X\times Y} [\langle \Lambda^*\overline{y}^*, x\rangle - \langle \overline{y}^*, y\rangle - J(x,y)] \\ &= -\sup_{(x,y)\in X\times Y} [\langle \Lambda^*\overline{y}^*, x\rangle - \langle \overline{y}^*, \Lambda x - y\rangle - J(x, \Lambda x - y)] \\ &= -\sup_{(x,y)\in X\times Y} [\langle \overline{y}^*, y\rangle - J(x, \Lambda x - y)] \\ &= \inf_{y\in Y} [h(y) - \langle \overline{y}^*, y\rangle] \\ &\geq \inf_{x\in X} F(x). \end{aligned}$$

By Lemma 2.2, it follows that $\overline{y}^*$ is a maximum of G and (2.1) holds. □

Corollary 2.5. *Suppose that (J,Λ) is suitable, and $(\underline{x}, \overline{y}^*) \in X \times Y^*$. Then $(\underline{x}, \overline{y}^*)$ satisfies the duality relation iff $\underline{x}$ is a minimum of F and $\overline{y}^*$ is a maximum of G.*

Proof. If $\underline{x}$ is a minimum of F and $\overline{y}^*$ is a maximum of G, then since (J,Λ) is suitable, the Fenchel–Rockafellar theorem furnishes a maximum $\tilde{y}^*$ of G such that

$$G(\overline{y}^*) = G(\tilde{y}^*) = \inf_{x\in X} F(x) = F(\underline{x}).$$

The converse is immediate from Lemma 2.2. □

3. The max flow/min cut theorem

Let (V,E) be a finite directed graph such that for any $(v,w) \in E$, $(w,v) \notin E$. The vector spaces $\mathbf{R}^V$ and $\mathbf{R}^E$ come with a natural inner product, which makes the Dirac delta functions at each vertex and edge form an orthonormal basis. The *exterior derivative* on (V,E) is

$$\begin{aligned} \mathrm{d} : \mathbf{R}^V &\to \mathbf{R}^E \\ f &\mapsto ((v,w) \mapsto f(w) - f(v)). \end{aligned}$$

Lemma 3.1 (divergence theorem)**.** *The adjoint $\mathrm{d}^* : \mathbf{R}^E \to \mathbf{R}^V$ of d satisfies*

$$\mathrm{d}^*\varphi(v) = \sum_{\substack{w\in V \\ (v,w)\in E}} \varphi(v,w) - \sum_{\substack{u\in V \\ (u,v)\in E}} \varphi(u,v).$$

Proof. We compute

$$\begin{aligned} \langle f, \mathrm{d}^*\varphi\rangle &= \langle \mathrm{d}f, \varphi\rangle \\ &= \sum_{(v,w)\in E} (f(w) - f(v))\varphi(v,w) \\ &\sum_{w\in V} \sum_{\substack{v\in V \\ (v,w)\in E}} f(w)\varphi(v,w) - \sum_{v\in V} \sum_{\substack{w\in V \\ (v,w)\in E}} f(v)\varphi(v,w) \\ &= \sum_{v\in V} f(v) \left[\sum_{\substack{w\in V \\ (v,w)\in E}} \varphi(v,w) - \sum_{\substack{u\in V \\ (u,v)\in E}} \varphi(u,v) \right]. \end{aligned}$$ □

Given two vertices $s_0, s_1 \in V$, let

$$J : \mathbf{R}^V \times \mathbf{R}^E \to \mathbf{R}$$

satisfy

$$J(f,\varphi) = \sum_{(v,w)\in E} |\varphi(v,w)|$$

if $f(s_i) = i$, and $J(f,\varphi) = +\infty$ otherwise. Then J is convex, and we are interested in minimizers $\underline{f}$ of $f \mapsto J(f, \mathrm{d}f)$, such a minimizer is called a *minimal cut*.

We compute the convex conjugate,

$$J^*(g,\psi) = \sup_{(f,\varphi)\in\mathbf{R}^V\times\mathbf{R}^E} \sum_{v\in V} f(v)g(v) + \sum_{(v,w)\in E} \varphi(v,w)\psi(v,w) - J(f,\varphi).$$

This supremum can only be attained by f such that $f(s_i) = i$, for otherwise J^* is identically $-\infty$. Thus

$$J^*(g,\psi) = \sup_{\substack{f\in\mathbf{R}^V\\ f(s_i)=i}} \sum_{v\in V} f(v)g(v) + \sup_{\varphi\in\mathbf{R}^E} \sum_{(v,w)\in E} (\varphi(v,w)\psi(v,w) - |\varphi(v,w)|).$$

The dual problem, thus, is to maximize

$$\begin{aligned} -J^*(\mathrm{d}^*\psi, -\psi) &= -\sup_{\substack{f\in\mathbf{R}^V\\ f(s_i)=i}} \sum_{v\in V} f(v)\,\mathrm{d}^*\psi(v) - \sup_{\varphi\in\mathbf{R}^E} \sum_{(v,w)\in E} (|\varphi(v,w)| - \varphi(v,w)\psi(v,w)) \\ &= \inf_{\substack{f\in\mathbf{R}^V\\ f(s_i)=-i}} \left[\sum_{v\in V} f(v)\,\mathrm{d}^*\psi(v)\right] + \inf_{\varphi\in\mathbf{R}^E} \sum_{(v,w)\in E} (\varphi(v,w)\psi(v,w) - |\varphi(v,w)|). \end{aligned}$$

Let V° be the interior, $V \setminus \{s_0, s_1\}$. In order for the first infimum to not just be $-\infty$, it must be that $\mathrm{d}^*\psi \restriction V^\circ = 0$, in which case the first infimum is just $-\mathrm{d}^*\psi(s_1)$. For the second infimum to not just be $-\infty$, it must be that $|\psi| \le 1$. We call $\overline{\psi}$ which maximizes $-J^*(\mathrm{d}^*\psi, -\psi)$ a *maximal cut*.

Theorem 3.2 (max flow/min cut theorem)**.** *There exists a minimal cut* $\underline{f}$ *and a maximal flow* $\overline{\psi}$*, and for any such minimal cut and maximal flow,*

$$\sum_{(v,w)\in E} |\,\mathrm{d}\underline{f}(v,w)| = -\,\mathrm{d}^*\overline{\psi}(s_1).$$

Proof. First observe that the infimum in

$$-J^*(\mathrm{d}^*\psi, -\psi) = -\,\mathrm{d}^*\psi(s_1) + \inf_{\varphi\in\mathbf{R}^E} \sum_{(v,w)\in E} (\varphi(v,w)\psi(v,w) - |\varphi(v,w)|),$$

if it is not $-\infty$, must be realized by a φ which has the opposite sign as ψ (since $|\psi| \le 1$). Therefore that infimum is ≤ 0 and so

$$-J^*(\mathrm{d}^*\psi, -\psi) \le -\,\mathrm{d}^*\psi(s_1).$$

Let $f_0(s_0) = 0$ and for every $v \ne s_0$, $f_0(v) = 1$. Then $0 \le J(f_0, \mathrm{d}f_0) < +\infty$ and $\psi \mapsto J(f_0, \psi)$ is continuous. So f_0 witnesses that (J, Λ) is suitable. Also $f \mapsto J(f, \mathrm{d}f)$ is coercive, since if $f(s_0) = 0$ and there exists $v \in V$ such that $|f(v)| \ge C$, then there exists $(u,w) \in E$ such that

$$|\,\mathrm{d}f(u,w)| \ge \frac{C}{\operatorname{card} E}.$$

So there is a pair $(\underline{f}, \overline{\psi})$ which satisfies the duality relation,

$$\sum_{(v,w)\in E} |\,\mathrm{d}\underline{f}(v,w)| = -\,\mathrm{d}^*\overline{\psi}(s_1) + \inf_{\varphi\in\mathbf{R}^E} \sum_{(v,w)\in E} (\varphi(v,w)\overline{\psi}(v,w) - |\varphi(v,w)|),$$

such that $\underline{f}(s_i) = i$, $|\overline{\psi}| \leq 1$, and $\mathrm{d}^* \overline{\psi} \restriction V^\circ = 0$. From the constraints $\underline{f}(s_i) = i$ and $\mathrm{d}^* \overline{\psi} \restriction V^\circ = 0$, we see that

$$-\mathrm{d}^* \overline{\psi}(s_1) = \sum_{v \in V} \underline{f}(v) \, \mathrm{d}^* \overline{\psi}(v),$$

and so

$$\begin{aligned}\sum_{(v,w) \in E} |\, \mathrm{d}\underline{f}(v,w)| &= \sum_{v \in V} \underline{f}(v) \, \mathrm{d}^* \overline{\psi}(v) + \inf_{\varphi \in \mathbf{R}^E} \sum_{(v,w) \in E} (\varphi(v,w)\overline{\psi}(v,w) - |\varphi(v,w)|) \\ &\leq \sum_{(v,w) \in V} \mathrm{d}\underline{f}(v,w)\overline{\psi}(v,w) + \sum_{(v,w) \in E} (\mathrm{d}\underline{f}(v,w)\overline{\psi}(v,w) - |\, \mathrm{d}\underline{f}(v,w)|).\end{aligned}$$

Rearranging and applying the constraint $|\overline{\psi}| \leq 1$,

$$\sum_{(v,w) \in E} |\, \mathrm{d}\underline{f}(v,w)| \leq \sum_{(v,w) \in E} \mathrm{d}\underline{f}(v,w)\overline{\psi}(v,w) \leq \sum_{(v,w) \in E} |\, \mathrm{d}\underline{f}(v,w)|.$$

The theorem follows, since

$$\sum_{(v,w) \in E} \mathrm{d}\underline{f}(v,w)\overline{\psi}(v,w) = -\mathrm{d}^* \psi(s_1). \qquad \square$$

Department of Mathematics, Brown University
Email address: aidan_backus@brown.edu