The conjugate function

- Definition and examples

Example 5 (Norm)**.** *Let* $\|\cdot\|$ *be a norm on* $\mathbb{R}^n$*, with dual norm* $\|\cdot\|_*$*. The conjugate of* $f(\mathbf{x}) = \|\mathbf{x}\|$ *is*

$$f^*(\mathbf{y}) = I_{\mathcal{B}}(\mathbf{y}),$$

where $\mathcal{B} = \{\mathbf{y} | \|\mathbf{y}\|_* \leq 1\}$ *is the unit ball of the dual norm.*

Example 6 (Norm squared)**.** *Now consider the function* $f(\mathbf{x}) = (1/2) \|\mathbf{x}\|^2$*, where* $\|\cdot\|$ *is a norm, with dual norm* $\|\cdot\|_*$*. Its conjugate is* $f^*(\mathbf{y}) = (1/2) \|\mathbf{y}\|_*^2$*.*

Example 7 (Support function)**.** $f(\mathbf{x}) = S_{\mathcal{C}}(\mathbf{x}) = \sup_{\mathbf{y} \in \mathcal{C}} \langle \mathbf{x}, \mathbf{y} \rangle$*, where* $\mathcal{C}$ *is a close convex set. Then its conjugate is the indicator function of* $\mathcal{C}$*:* $f^*(\mathbf{y}) = I_{\mathcal{C}}(\mathbf{y})$*.*

The conjugate function

- Basic properties

Fenchel's inequality:

$$f(\mathbf{x}) + f^*(\mathbf{y}) \geq \langle \mathbf{x}, \mathbf{y} \rangle, \quad \forall \mathbf{x},\ \mathbf{y}.$$

For example, if $f(\mathbf{x}) = \frac{1}{2}\|\mathbf{x}\|^2$ then $f^*(\mathbf{y}) = \frac{1}{2}\|\mathbf{y}\|_*^2$ and thus

$$\frac{1}{2}\|\mathbf{x}\|^2 + \frac{1}{2}\|\mathbf{y}\|_*^2 \geq \langle \mathbf{x}, \mathbf{y} \rangle,$$

which is generalization of $\frac{1}{2}\|\mathbf{x}\|_2^2 + \frac{1}{2}\|\mathbf{y}\|_2^2 \geq \langle \mathbf{x}, \mathbf{y} \rangle$.

When $f(\mathbf{x}) = (1/2)\mathbf{x}^T\mathbf{Q}\mathbf{x}$, where $\mathbf{Q} \in \mathbb{S}_{++}^n$, we obtain the inequality

$$\langle \mathbf{x}, \mathbf{y} \rangle \leq (1/2)\mathbf{x}^T\mathbf{Q}\mathbf{x} + (1/2)\mathbf{y}^T\mathbf{Q}^{-1}\mathbf{y}.$$

The conjugate function

- Basic properties

Conjugate of the conjugate

Theorem: If f is proper, convex and closed, then $f^{**} = f$.

Proof. It is easy to prove that $f^{**} \leq f$, i.e. $\text{epi} f^{**} \subseteq \text{epi} f$. Suppose $\exists (\mathbf{x}_0^T, \gamma)^T \notin \text{epi} f$, where $\gamma \geq f^{**}(\mathbf{x}_0)$. Since f is proper, convex and closed, its epigraph is closed and does not include a vertical line. Thus there exists a hyperplane parameterized as $(\mathbf{w}^T, \zeta)^T$ $(\zeta \neq 0)$ and μ such that it strictly separates $\text{epi} f$ and the point $(\mathbf{x}_0^T, \gamma)^T$:

$$(\mathbf{w}^T, \zeta)(\mathbf{x}^T, t)^T < \mu < (\mathbf{w}^T, \zeta)(\mathbf{x}_0^T, \gamma)^T, \quad \forall \mathbf{x} \in \text{dom}\, f, t \geq f(\mathbf{x}).$$

Since t can be arbitrarily large, ζ must be negative. W.l.o.g, we assume $\zeta = -1$. Then taking $t = f(\mathbf{x})$,

$$\mathbf{w}^T \mathbf{x} - f(\mathbf{x}) < \mu < \mathbf{w}^T \mathbf{x}_0 - \gamma \leq \mathbf{w}^T \mathbf{x}_0 - f^{**}(\mathbf{x}_0), \quad \forall \mathbf{x} \in \text{dom}\, f.$$

Thus $f^*(\mathbf{w}) < \mathbf{w}^T \mathbf{x}_0 - f^{**}(\mathbf{x}_0)$, contradicting Fenchel's inequality.

The conjugate function

- Basic properties

Useful for finding a convex surrogate of a nonconvex function!

Conjugate of the conjugate

Theorem: For any f, f^{**} is the largest convex function not exceeding f, i.e., *convex envelop* of f.

Proof. We first have $f^{***}(\mathbf{y}) = f^{*}(\mathbf{y})$ because $f^{*}(\mathbf{y})$ is a propoer, convex and closed function. And it is easy to prove that if $f_1(\mathbf{x}) \leq f_2(\mathbf{x})$, then $f_1^{*}(\mathbf{y}) \geq f_2^{*}(\mathbf{y})$.

Suppose g is proper, convex and closed, and $f^{**}(\mathbf{x}) \leq g(\mathbf{x}) \leq f(\mathbf{x})$. Then by the above results,

$$f^{*}(\mathbf{y}) \leq g^{*}(\mathbf{y}) \leq f^{***}(\mathbf{y}) = f^{*}(\mathbf{y}).$$

Thus $f^{*}(\mathbf{y}) = g^{*}(\mathbf{y})$ and hence $f^{**} = g^{**} = g$.

The conjugate function

- Basic properties

Conjugate of the conjugate

Examples. 1. The convex envelope of ℓ_0-pseudo-norm on unit ℓ_1 ball is the ℓ_1-norm.

Proof. We first compute $f^*(\mathbf{y}) = \sup_{\|\mathbf{x}\| \le 1}(\langle \mathbf{y}, \mathbf{x} \rangle - \|\mathbf{x}\|_0)$. When $\|\mathbf{x}\| \le 1$,

$$f^*(\mathbf{y}) \le \|\mathbf{y}\|_\infty \|\mathbf{x}\|_1 - \|\mathbf{x}\|_0 \le \|\mathbf{y}\|_\infty - \|\mathbf{x}\|_0.$$

$$f^*(\mathbf{y}) = \begin{cases} \|\mathbf{y}\|_\infty - 1, & \text{if } \|\mathbf{y}\|_\infty \ge 1, \\ 0, & \text{if } \|\mathbf{y}\|_\infty < 1 \end{cases} = \max(\|\mathbf{y}\|_\infty - 1, 0).$$

Next, we compute $f^{**}(\mathbf{x}) = \sup_{\mathbf{y}}(\langle \mathbf{y}, \mathbf{x} \rangle - f^*(\mathbf{y}))$:

$$f^{**}(\mathbf{x}) \le \|\mathbf{y}\|_\infty \|\mathbf{x}\|_1 - \max(\|\mathbf{y}\|_\infty - 1, 0).$$

$$f^{**}(\mathbf{x}) = \begin{cases} \|\mathbf{x}\|_1, & \text{if } \|\mathbf{x}\|_1 \le 1, \\ +\infty, & \text{if } \|\mathbf{x}\|_1 > 1. \end{cases}$$

The conjugate function

- Basic properties

Conjugate of the conjugate

Examples. 2. The convex envelope of rank on unit 2-norm ball is the nuclear-norm.

Proof. 1) Compute rank^*. Let $q = \min(m, n)$. By von Neumann's trace thm,

$$\mathrm{rank}^*(\mathbf{Y}) = \sup_{\|\mathbf{X}\|_2 \leq 1} \{\langle \mathbf{X}, \mathbf{Y} \rangle - \mathrm{rank}(\mathbf{X})\} \leq \sup_{\|\mathbf{X}\|_2 \leq 1} \left\{ \sum_{i=1}^{q} \sigma_i(\mathbf{X}) \sigma_i(\mathbf{Y}) - \mathrm{rank}(\mathbf{X}) \right\}.$$

The equality can hold when $\mathbf{X}$ and $\mathbf{Y}$ have the same singular subspaces. Suppose $\mathrm{rank}(\mathbf{X}) = r$, $0 \leq r \leq q$. Then

$$\mathrm{rank}^*(\mathbf{Y}) = \sup_{\|\mathbf{X}\|_2 \leq 1} \left\{ \sum_{i=1}^{r} \sigma_i(\mathbf{Y}) - r \right\} = \sup_{\|\mathbf{X}\|_2 \leq 1} \left\{ \sum_{i=1}^{r} (\sigma_i(\mathbf{Y}) - 1) \right\}.$$

So r should be chosen such that $\sigma_r(\mathbf{Y}) > 1$ and $\sigma_{r+1}(\mathbf{Y}) \leq 1$, and

$$\mathrm{rank}^*(\mathbf{Y}) = \sum_{i=1}^{r} (\sigma_i(\mathbf{Y}) - 1).$$

The conjugate function

- Basic properties

2) Compute rank^{**}. By von Neumann's trace thm. again,

$$\text{rank}^{**}(\mathbf{X}) = \sup_{\mathbf{Y}}\{\langle \mathbf{X}, \mathbf{Y}\rangle - \text{rank}^*(\mathbf{Y})\} \le \sup_{\mathbf{Y}} \left\{ \sum_{i=1}^{q} \sigma_i(\mathbf{X})\sigma_i(\mathbf{Y}) - \sum_{i=1}^{r} (\sigma_i(\mathbf{Y}) - 1) \right\},$$

where r is such that $\sigma_r(\mathbf{Y}) > 1$ and $\sigma_{r+1}(\mathbf{Y}) \le 1$. The equality can also hold.

If $\|\mathbf{X}\|_2 > 1$, we can choose $\sigma_1(\mathbf{Y})$ large enough so that $\text{rank}^{**}(\mathbf{X}) \to +\infty$, because the coefficient of $\sigma_1(\mathbf{Y})$ is a positive value $\sigma_1(\mathbf{X}) - 1$. If $\|\mathbf{X}\|_2 \le 1$,

$$\begin{aligned}
&\sum_{i=1}^{q} \sigma_i(\mathbf{X})\sigma_i(\mathbf{Y}) - \sum_{i=1}^{r} (\sigma_i(\mathbf{Y}) - 1) \\
&= \sum_{i=1}^{q} \sigma_i(\mathbf{X})\sigma_i(\mathbf{Y}) - \sum_{i=1}^{r} (\sigma_i(\mathbf{Y}) - 1) - \sum_{i=1}^{q} \sigma_i(\mathbf{X}) + \sum_{i=1}^{q} \sigma_i(\mathbf{X}) \\
&= \sum_{i=1}^{r} (\sigma_i(\mathbf{X}) - 1)(\sigma_i(\mathbf{Y}) - 1) + \sum_{i=r+1}^{q} \sigma_i(\mathbf{X})(\sigma_i(\mathbf{Y}) - 1) + \sum_{i=1}^{q} \sigma_i(\mathbf{X}) \\
&\le \sum_{i=1}^{q} \sigma_i(\mathbf{X}) = \|\mathbf{X}\|_*.
\end{aligned}$$

The equality can hold. Thus $\text{rank}^{**}(\mathbf{X}) = \|\mathbf{X}\|_*$ over the set $\{\mathbf{X} | \|\mathbf{X}\|_2 \le 1\}$.

The conjugate function

- Basic properties

Useful for finding the gradient for the dual problem!

Differentiable functions

The conjugate of a differentiable function f is also called the *Legendre transform* of f.

Suppose f is convex and differentiable, with $\boxed{\operatorname{dom} f = \mathbf{R}^n}$. Any maximizer $\mathbf{x}^*$ of $\langle \mathbf{y}, \mathbf{x} \rangle - f(\mathbf{x})$ satisfies $\mathbf{y} = \nabla f(\mathbf{x}^*)$, and conversely, if $\mathbf{x}^*$ satisfies $\mathbf{y} = \nabla f(\mathbf{x}^*)$, then $\mathbf{x}^*$ maximizes $\langle \mathbf{y}, \mathbf{x} \rangle - f(\mathbf{x})$. Therefore, if $\mathbf{y} = \nabla f(\mathbf{x}^*)$, we have

$$f^*(\mathbf{y}) = \mathbf{x}^{*T} \nabla f(\mathbf{x}^*) - f(\mathbf{x}^*).$$

This allows us to determine $f^*(\mathbf{y})$ for any $\mathbf{y}$ for which we can solve the gradient equation $\mathbf{y} = \nabla f(\mathbf{z})$ for $\mathbf{z}$.

We can express this another way. Let $\mathbf{z} \in \mathbf{R}^n$ be arbitrary and define $\mathbf{y} = \nabla f(\mathbf{z})$. Then we have

$$f^*(\mathbf{y}) = \mathbf{z}^T \nabla f(\mathbf{z}) - f(\mathbf{z}).$$

If f is convex and closed, then

$$\mathbf{y} \in \partial f(\mathbf{x}) \Leftrightarrow \mathbf{x} \in \partial f^*(\mathbf{y}) \Leftrightarrow f(\mathbf{x}) + f^*(\mathbf{y}) = \langle \mathbf{x}, \mathbf{y} \rangle.$$

The conjugate function

- Basic properties

Scaling and composition with affine transformation

For $a > 0$ and $b \in \mathbb{R}$, the conjugate of $g(x) = af(x) + b$ is $g^*(y) = af^*(y/a) - b$.

Suppose $\mathbf{A} \in \mathbb{R}^{n \times n}$ is nonsingular and $\mathbf{b} \in \mathbb{R}^n$. Then the conjugate of $g(\mathbf{x}) = f(\mathbf{A}\mathbf{x} + \mathbf{b})$ is

$$g^*(\mathbf{y}) = f^*(\mathbf{A}^{-T}\mathbf{y}) - \mathbf{b}^T\mathbf{A}^{-T}\mathbf{y},$$

with $\operatorname{dom} g^* = \mathbf{A}^T \operatorname{dom} f^*$.

The conjugate function

- Basic properties

Separable functions

If $f(\mathbf{u}, \mathbf{v}) = f_1(\mathbf{u}) + f_2(\mathbf{v})$, where f_1 and f_2 are convex functions with conjugates f_1^* and f_2^*, respectively, then

$$f^*(\mathbf{w}, \mathbf{z}) = f_1^*(\mathbf{w}) + f_2^*(\mathbf{z}).$$

In other words, the conjugate of a *separable* convex function is the sum of the conjugates.

Envelope function and Proximal mapping

Let $f : \mathbb{R}^n \to (-\infty, +\infty]$ be a closed proper function. For a scalar $c > 0$, define the corresponding *envelope function* $\mathrm{Env}_c f$ and the *proximal mapping* $\mathrm{Prox}_c\, f$ by

Involution: $(f \boxtimes g)(\mathbf{x}) = \inf_{\mathbf{y}+\mathbf{z}=\mathbf{x}} f(\mathbf{y}) + g(\mathbf{z})$.

$$\begin{aligned} \mathrm{Env}_c f(\mathbf{x}) &= \inf_{\mathbf{w}} \left\{ f(\mathbf{w}) + \frac{1}{2c}\|\mathbf{w}-\mathbf{x}\|^2 \right\}, \\ \mathrm{Prox}_c\, f(\mathbf{x}) &= \operatorname*{argmin}_{\mathbf{w}} \left\{ f(\mathbf{w}) + \frac{1}{2c}\|\mathbf{w}-\mathbf{x}\|^2 \right\}. \end{aligned} \tag{1}$$

Extremely useful for updating iterates!

Remark:

1. The envelope function $\mathrm{Env}_c f$ is an underestimate of the function f, i.e., $\mathrm{Env}_c f(\mathbf{x}) \le f(\mathbf{x})$, $\forall \mathbf{x} \in \mathbb{R}^n$. Furthermore, $\mathrm{Env}_c f$ is a real-valued continuous function, whereas f itself may only be extended real-valued and lower semi-continuous.

2. It is obvious that $\mathbf{u} = \mathrm{Prox}_c\, f(\mathbf{x}) \Leftrightarrow \mathbf{x} - \mathbf{u} \in \partial f(\mathbf{u})$.

Envelope function and Proximal mapping

Examples: $f(\mathbf{x}) = \chi_C(\mathbf{x})$

Projection onto halfspace: $\mathcal{C} = \{\mathbf{x} | \mathbf{a}^T\mathbf{x} \le b\}$ ($\mathbf{a} \neq \mathbf{0}$):

$$P_{\mathcal{C}}(\mathbf{x}) = \begin{cases} \mathbf{x} + \frac{b - \mathbf{a}^T\mathbf{x}}{\|\mathbf{a}\|^2}\mathbf{a}, & \text{if } \mathbf{a}^T\mathbf{x} > b, \\ \mathbf{x}, & \text{if } \mathbf{a}^T\mathbf{x} \le b. \end{cases}$$

Projection onto hyperbox $\mathcal{C} = \{\mathbf{x} | \mathbf{l} \le \mathbf{x} \le \mathbf{u}\}$:

$$(P_{\mathcal{C}}(\mathbf{x}))_i = \begin{cases} l_i, & \text{if } x_i \le l_i, \\ x_i, & \text{if } l_i \le \mathbf{x}_i \le u_i, \\ u_i, & \text{if } x_i \ge u_i. \end{cases}$$

Projection onto nonnegative orthant: $\mathcal{C} = \mathbb{R}^n_+$:

$$P_{\mathcal{C}}(\mathbf{x}) = \max(\mathbf{x}, \mathbf{0}).$$

Envelope function and Proximal mapping

Projection onto positive semi-definite cone: $\mathcal{C} = \mathbb{S}^n_+$:

$$P_{\mathcal{C}}(\mathbf{X}) = \sum_{i=1}^{n} \max(\lambda_i, 0)\mathbf{q}_i\mathbf{q}_i^T,$$

where $\sum_{i=1}^{n} \lambda_i \mathbf{q}_i \mathbf{q}_i^T$ is the eigenvalue decomposition of $\mathbf{X}$.
Projection onto unit Euclidean ball: $\mathcal{C} = \{\mathbf{x} | \|\mathbf{x}\| \leq 1\}$:

$$P_{\mathcal{C}}(\mathbf{x}) = \begin{cases} \frac{1}{\|\mathbf{x}\|}\mathbf{x}, & \text{if } \|\mathbf{x}\| > 1, \\ \mathbf{x}, & \text{if } \|\mathbf{x}\| \leq 1. \end{cases}$$

Projection onto unit matrix 2-norm ball: $\mathcal{C} = \{\mathbf{X} | \|\mathbf{X}\|_2 \leq 1\}$:

$$P_{\mathcal{C}}(\mathbf{X}) = \begin{cases} \mathbf{U}\,\mathrm{Diag}(\min(\boldsymbol{\sigma}, 1))\mathbf{V}^T, & \text{if } \|\mathbf{X}\|_2 > 1, \\ \mathbf{X}, & \text{if } \|\mathbf{X}\|_2 \leq 1. \end{cases}$$

where $\mathbf{U}\,\mathrm{Diag}(\boldsymbol{\sigma})\mathbf{V}^T$ is the SVD of $\mathbf{X}$.

Envelope function and Proximal mapping

Examples: $f(\mathbf{x}) = \frac{1}{2}\mathbf{x}^T\mathbf{A}\mathbf{x} + \mathbf{b}^t\mathbf{x} + \mathbf{c}$, $\|\mathbf{x}\|_1$, $\sum_{i=1}^{n} \log x_i$, $f(\mathbf{X}) = \|\mathbf{X}\|_*$

Theorem 1. *For each $\tau > 0$ and $\mathbf{Y} \in \mathbb{R}^{n_1 \times n_2}$,*

$$\mathcal{D}_\tau(\mathbf{Y}) = \underset{\mathbf{X}}{\operatorname{argmin}} \left\{ \tau\|\mathbf{X}\|_* + \frac{1}{2}\|\mathbf{X} - \mathbf{Y}\|_F^2 \right\}, \tag{1}$$

where $\mathcal{D}_\tau(\mathbf{Y})$ is the singular value thresholding operator defined as

$$\mathcal{D}_\tau(\mathbf{Y}) = \mathbf{U}\operatorname{diag}(\{(\sigma_i - \tau)_+\})\mathbf{V}^T, \tag{2}$$

in which $\mathbf{U}\operatorname{diag}(\{\sigma_i\})\mathbf{V}^T$ is the SVD of $\mathbf{Y}$.

Envelope function and Proximal mapping

Separable functions
If $f(\mathbf{u}, \mathbf{v}) = f_1(\mathbf{u}) + f_2(\mathbf{v})$, where f_1 and f_2 are closed proper functions, then

$$\mathrm{Prox}_c f(\mathbf{u}, \mathbf{v}) = \mathrm{Prox}_c f_1(\mathbf{u}) + \mathrm{Prox}_c f_2(\mathbf{v}).$$

Scaling and composition with orthogonal transformation
For $a > 0$ and $b \in \mathbb{R}$, the conjugate of $g(x) = f(ax + b)$ is

$$\mathrm{Prox}_c g(x) = a^{-1}(\mathrm{Prox}_{ca^2} f(ax + b) - b).$$

Suppose $g(\mathbf{x}) = f(\mathbf{A}\mathbf{x} + \mathbf{b})$, where $\mathbf{A} \in \mathbb{R}^{n \times m}$ satisfies $\mathbf{A}\mathbf{A}^T = \lambda^{-1}\mathbf{I}$ $(\lambda > 0)$ and $\mathbf{b} \in \mathbb{R}^n$. Then

$$\mathrm{Prox}_c g(\mathbf{x}) = (\mathbf{I} - \lambda \mathbf{A}^T \mathbf{A})\mathbf{x} + \lambda \mathbf{A}^T (\mathrm{Prox}_{c\lambda^{-1}} f(\mathbf{A}\mathbf{x} + \mathbf{b}) - \mathbf{b}).$$

Envelope function and Proximal mapping

Proof. $\mathbf{w} = \mathrm{Prox}_c\, g(\mathbf{x})$ is the solution of the optimization problem:

$$\min_{\mathbf{w},\mathbf{z}} f(\mathbf{z}) + \frac{1}{2c}\|\mathbf{w}-\mathbf{x}\|^2, \quad s.t. \quad \mathbf{A}\mathbf{w}+\mathbf{b}=\mathbf{z}.$$

Eliminating $\mathbf{w}$ gives: $\mathbf{w} = \mathbf{x} + \mathbf{A}^T(\mathbf{A}\mathbf{A}^T)^{-1}(\mathbf{z}-\mathbf{b}-\mathbf{A}\mathbf{x}) = (\mathbf{I}-\lambda\mathbf{A}^T\mathbf{A})\mathbf{x} + \lambda\mathbf{A}^T(\mathbf{z}-\mathbf{b})$. The optimal $\mathbf{z}$ is the minimizer of

$$f(\mathbf{z}) + \frac{\lambda^2}{2c}\|\mathbf{A}^T(\mathbf{z}-\mathbf{b}-\mathbf{A}\mathbf{x})\|^2 = f(\mathbf{z}) + \frac{\lambda}{2c}\|\mathbf{z}-\mathbf{b}-\mathbf{A}\mathbf{x}\|^2,$$

which is $\mathbf{z} = \mathrm{Prox}_{c\lambda^{-1}}\, f(\mathbf{A}\mathbf{x}+\mathbf{b})$.

Envelope function and Proximal mapping

Theorem 2. *For any function f, $\operatorname{Prox}_c f(\mathbf{x})$ is a monotonic function of $\mathbf{x}$ in the sense that:*

$$\langle \mathbf{x}_1 - \mathbf{x}_2, \mathbf{y}_1 - \mathbf{y}_2 \rangle \geq 0, \quad \forall \mathbf{y}_i \in Prox_c f(\mathbf{x}_i), i = 1, 2. \tag{1}$$

Proposition 1. *Let $f : \mathbb{R}^n \to (-\infty, +\infty]$ be a closed proper convex function and $c > 0$. The envelope function $Env_c f$ is convex and smooth and its gradient is given by*

$$\nabla Env_c f(\mathbf{x}) = \frac{1}{c}(\mathbf{x} - Prox_c f(\mathbf{x})). \tag{2}$$

The envelope function $\mathrm{Env}_c f$ is smooth, regardless of whether f is smooth.

Envelope function and Proximal mapping

Proposition 2. *Let* $f : \mathbb{R}^n \to (-\infty, +\infty]$ *be a closed proper convex function and* $c > 0$*. The proximal mapping* $\operatorname{Prox}_c f$ *is single-valued and is continuous:* $\operatorname{Prox}_c f(x) \to \operatorname{Prox}_{c^*} f(\mathbf{x}^*)$ *whenever* $(\mathbf{x}, c) \to (\mathbf{x}^*, c^*)$*, with* $c^* > 0$*.*

Theorem 3 (Moreau Decomposition)**.** $\mathbf{x} = \operatorname{Prox}_c f(\mathbf{x}) + c \operatorname{Prox}_{c^{-1}} f^*(c^{-1}\mathbf{x})$.

Proof. Let $\mathbf{u} = \operatorname{Prox}_c f(\mathbf{x})$. Then

$$\begin{aligned} \mathbf{u} = \operatorname{Prox}_c f(\mathbf{x}) &\iff -c^{-1}(\mathbf{u} - \mathbf{x}) \in \partial f(\mathbf{u}) \\ &\iff \mathbf{u} \in \partial f^*(-c^{-1}(\mathbf{u} - \mathbf{x})) \end{aligned}$$

Let $\mathbf{z} = -c^{-1}(\mathbf{u} - \mathbf{x})$. Then $\mathbf{u} = \mathbf{x} - c\mathbf{z}$ and thus

$$\begin{aligned} &\iff \mathbf{x} - c\mathbf{z} \in \partial f^*(\mathbf{z}) \\ &\iff \mathbf{0} \in \partial f^*(\mathbf{z}) + c(\mathbf{z} - c^{-1}\mathbf{x}) \\ &\iff \mathbf{z} = \operatorname{Prox}_{c^{-1}} f^*(c^{-1}\mathbf{x}). \end{aligned}$$

So $\mathbf{x} = \mathbf{u} + c\mathbf{z} = \operatorname{Prox}_c f(\mathbf{x}) + c \operatorname{Prox}_{c^{-1}} f^*(c^{-1}\mathbf{x})$.

Envelope function and Proximal mapping

Example: proximal mapping of a norm.
We know that if $f(\mathbf{x}) = \|\mathbf{x}\|$, then $f^*(\mathbf{y}) = I_{\mathcal{B}}(\mathbf{y})$, where $\mathcal{B}$ is the unit ball of the dual norm $\|\cdot\|_*$. Then by Moreay decomposition:

$$\begin{aligned}
\mathrm{Prox}_c f(\mathbf{x}) &= \mathbf{x} - c\mathrm{Prox}_{c^{-1}} f^*(\mathbf{x}/c) \\
&= \mathbf{x} - cP_{\mathcal{B}}(\mathbf{x}/c) \\
&= \mathbf{x} - P_{c\mathcal{B}}(\mathbf{x}).
\end{aligned}$$

Examples: $f(\mathbf{x}) = \|\mathbf{x}\|_1$, $\|\mathbf{x}\|_2$, $f(\mathbf{X}) = \|\mathbf{X}\|_2$, $\|\mathbf{X}\|_*$.