Cones Reference

Theo Diamandis Chandler Squires

December 18, 2020

1 Projections onto Convex Cones and their Derivatives

To compute projections and derivatives of projections onto the dual cone, we use the Moreau decomposition:

$$x = \Pi_{\mathcal{K}}(x) + \Pi_{\mathcal{K}^o}(x), \tag{1}$$

where $\mathcal{K}^o = -\mathcal{K}^\star$ is the polar cone. Thus,

$$\Pi_{\mathcal{K}^\star}(x) = -\Pi_{\mathcal{K}^o}(-x) = x + \Pi_{\mathcal{K}}(-x). \tag{2}$$

Differentiating, we obtain

$$D\Pi_{\mathcal{K}^\star}(x) = I - D\Pi_{\mathcal{K}}(-x). \tag{3}$$

Equipped with these relations, we will only consider projections onto primal cones in this section.

1.1 Zero Cone

Let $\mathcal{K} = \{0\}^n$, dual to $\mathcal{K}^\star = \mathbf{R}^n$. Then

$$\Pi_{\mathcal{K}}(x) = 0, \qquad D\Pi_{\mathcal{K}}(x) = 0, \tag{4}$$

where 0 is the appropriately-sized zero vector and zero matrix respectively.

1.2 Positive Orthant

Let $\mathcal{K} = \{x \mid x \succeq 0\}$, which is self-dual. Then

$$\Pi_{\mathcal{K}}(x) = \max(x, 0), \qquad D\Pi_{\mathcal{K}}(x) = \frac{1}{2}\,\mathbf{diag}\,(\mathrm{sign}(x) + \mathbf{1}), \tag{5}$$

where max and sign are taken elementwise.

1.3 Second-order Cone

Let $\mathcal{K} = \{(t, x) \mid \|x\|_2 \le t\}$, which is self-dual. The projection onto the second-order cone is given by [3, Ch 8]

$$\Pi_{\mathcal{K}}(x) = \begin{cases} 0 & \|x\| \le -t \quad (x \in \mathcal{K}^o) \\ (t, x) & \|x\| \le t \quad (x \in \mathcal{K}) \\ \frac{1}{2}(1 + t/\|x\|)(\|x\|, x) & \|x\| \ge |t|. \end{cases} \tag{6}$$

The derivative is

$$D\Pi_{\mathcal{K}}(x) = \begin{cases} 0 & \|x\| \leq -t \\ I & \|x\| \leq t \\ D & \|x\| \geq |t|, \end{cases} \qquad \text{where} \qquad D = \frac{1}{2\|x\|}\begin{bmatrix} \|x\| & x^T \\ x & (\|x\|+t)I - \frac{t}{\|x\|^2}xx^T \end{bmatrix}$$

1.4 Positive Semidefinite Cone

Let $\mathcal{K} = \{X \in \mathbf{S}^n_+ \mid \lambda_{\min}(X) \geq 0\}$, which is self-dual. Then

$$\Pi_{\mathcal{K}}(X) = \sum_{i=1}^{t} \max(\lambda_i, 0) v_i v_i^T, \tag{7}$$

where $X = \sum_{i=1}^n \lambda_i v_i v_i^T = V\Lambda V^T$ is the eigenvalue decomposition of X [3, Ch 8]. The derivative applied to a perturbation $\mathtt{d}X$ is [4]

$$D\Pi_{\mathcal{K}}(X)(\mathtt{d}X) = V(B \circ (V^T \mathtt{d}X V))V^T, \qquad \text{where} \qquad [B]_{i,j} = \begin{cases} 0 & i \leq k,\ j \leq k \\ \frac{(\lambda_i)_+}{(\lambda_j)_- + (\lambda_i)_+} & i > k,\ j \leq k \\ \frac{(\lambda_j)_+}{(\lambda_i)_- + (\lambda_j)_+} & i \leq k,\ j > k \\ 1 & i > k,\ j > k, \end{cases} \tag{8}$$

where $\circ$ is the Hadamard product, k is the number of negative eigenvalues of X, $(x)_+ = \max(x, 0)$, and $(x)_- = \max(-x, 0)$. Note that we can write $D\Pi_{\mathcal{K}}(X)$ as a matrix that operates on $\tilde{x} = \mathbf{vec}(X)$, which we use `ConeProgramDiff.jl`.

1.5 Exponential Cone

Let $\mathcal{K} = \{(x, y, z) \mid ye^{x/y} \leq z,\ y > 0\}$, which is dual to $\mathcal{K}^\star = \{(u, v, w) \mid -ue^{v/u} \leq ew,\ u < 0\}$. The projection has four cases [6, Ch 6.3.4]

$$\Pi_{\mathcal{K}}(x, y, z) = \begin{cases} (x, y, z) & (x, y, z) \in \mathcal{K} \\ (0, 0, 0) & (x, y, z) \in \mathcal{K}^o \\ (x, y, \max(z, 0)) & x < 0,\ y < 0 \\ \operatorname{argmin}_{(\tilde{x}, \tilde{y}, \tilde{z})} \|(x, y, z) - (\tilde{x}, \tilde{y}, \tilde{z})\|^2 \text{ s.t } \tilde{z} = \tilde{y}e^{\tilde{x}/\tilde{y}},\ \tilde{y} > 0 & \text{otherwise.} \end{cases} \tag{9}$$

The last problem can be solved via a primal-dual Newton method, as suggested in [6], or by an equivalent one-dimensional root finding problem, which is the approach used in the MOSEK solver [2] and outlined in [7]. In the fourth case

$$\Pi_{\mathcal{K}}(x, y, z) = \frac{(\rho - 1)x + y}{\rho^2 - \rho + 1}(e^\rho, 1, \rho), \tag{10}$$

where ρ is the solution to

$$\frac{((\rho - 1)x + y)e^\rho - (x - \rho y)e^{-\rho}}{\rho^2 - \rho + 1} - z = 0. \tag{11}$$

We can start with upper bound $u = r/y$ if $y > 0$ or lower bound $l = 1 - y/x$. if $x > 0$. The derivative similarly requires four cases [4, 1]:

$$D\Pi_{\mathcal{K}}(x,y,z) = \begin{cases} I & (x,y,z) \in \mathcal{K} \\ 0 & (x,y,z) \in \mathcal{K}^o \\ \mathbf{diag}\left((1,0,\max(0,z))\right) & x < 0,\ y < 0\ z \neq 0 \\ D & \text{otherwise.} \end{cases} \tag{12}$$

Let (r,s,t) be the solution to $\operatorname{argmin}_{(\tilde{x},\tilde{y},\tilde{z})} \|(x,y,z) - (\tilde{x},\tilde{y},\tilde{z})\|^2$ s.t $\tilde{z} = \tilde{y}e^{\tilde{x}/\tilde{y}}$, $\tilde{y} > 0$ and $\mu = t - z$ be the optimal dual variable associated with the equality constraint (from the KKT conditions). The KKT conditions can be implicitly differentiated through, and D is the upper left 3×3 block of the matrix

$$\begin{bmatrix} 1 + \frac{\mu e^{r/s}}{s} & -\frac{\mu r e^{r/s}}{s^2} & 0 & e^{r/s} \\ -\frac{\mu r e^{r/s}}{s^2} & 1 + \frac{\mu r^2 e^{r/s}}{s^3} & 0 & (1 - r/s)e^{r/s} \\ 0 & 0 & 1 & -1 \\ e^{r/s} & (1 - r/s)e^{r/s} - 1 & -1 & 0 \end{bmatrix}^{-1}, \tag{13}$$

which is invertible because the optimization problem is feasible.

1.6 Power Cone

Let $\mathcal{K} = \{(x,y,z) \mid x^\alpha y^{1-\alpha} \geq |z|,\ x \geq 0, y \geq 0\}$, which is dual to $\mathcal{K} = \{(u,v,w) \mid \left(\frac{u}{\alpha}\right)^\alpha \left(\frac{v}{1-\alpha}\right)^{1-\alpha} \geq |w|, \quad u \geq 0, \quad v \geq 0\}$ for $\alpha \in (0,1)$. The projection is found through the KKT conditions of the projection optimization problem, using a barrier function for the constraint and considering the limit. The derivative is found by the implicit function theorem. Of course, analysis requires significant calculation and application of some additional results, which can be found in [5]. The projection onto the power cone is given by [5, Prop. 2.2]:

$$\Pi_{\mathcal{K}}(x,y,z) = \begin{cases} (x,y,z) & (x,y,z) \in \mathcal{K} \\ (0,0,0) & (x,y,z) \in \mathcal{K}^o \\ (\max(x,0), \max(y,0), 0) & (x,y,z) \notin \mathcal{K}, \mathcal{K}^o,\ z = 0 \\ (\frac{1}{2}\phi(x,\alpha), \frac{1}{2}\phi(y,1-\alpha), \operatorname{sign}(z) \cdot r) & \text{otherwise,} \end{cases} \tag{14}$$

where $\phi(w,\beta) = w + \sqrt{w^2 + 4\beta r(|z| - r)}$ and r is the unique solution to the system

$$\frac{1}{2}\phi(x,\alpha)^\alpha \phi(y,1-\alpha)^{1-\alpha} - r = 0, \qquad 0 < r < |z|. \tag{15}$$

Let $I_+(t)$ be the indicator function of $\{t > 0\}$, $v = \begin{bmatrix} x & y \end{bmatrix}$, and $\bar{\alpha} = \begin{bmatrix} \alpha & 1-\alpha \end{bmatrix}$. The derivative of this projection is [5, Theorem 3.1]

$$D\Pi_{\mathcal{K}}(x,y,z) = \begin{cases} I & (x,y,z) \in \mathcal{K} \\ 0 & (x,y,z) \in \mathcal{K}^o \\ \mathbf{diag}\left((I_+(x), I_+(y), d)\right) & (x,y,z) \notin \mathcal{K}, \mathcal{K}^o,\ z = 0 \\ D & \text{otherwise,} \end{cases} \tag{16}$$

where

$$d = \begin{cases} 1 & \sum_{v_i>0} \bar{\alpha}_i > \sum_{v_i<0} \bar{\alpha}_i \\ 0 & \sum_{v_i>0} \bar{\alpha}_i < \sum_{v_i<0} \bar{\alpha}_i \\ \frac{1}{\left(\frac{\prod_{v_i<0}(-v_i)^{\bar{\alpha}_i}}{\prod_{v_i<0} \bar{\alpha}_i^{\bar{\alpha}_i} \prod_{v_i>0}(v_i)^{\bar{\alpha}_i}}\right)^2+1} & \sum_{v_i>0} \bar{\alpha}_i = \sum_{v_i<0} \bar{\alpha}_i, \end{cases} \tag{17}$$

and

$$D = \begin{bmatrix} \frac{1}{2} + \frac{x}{2g_x} + \frac{\alpha^2(|z|-2r)rL}{g_x^2} & -\alpha(1-\alpha)(|z|-2r)T & \text{sign}(z)\frac{\alpha * r * L}{g_x} \\ -\alpha(1-\alpha)(|z|-2r)T & \frac{1}{2} + \frac{y}{2g_y} + \frac{(1-\alpha)2(|z|-2r)rL}{g_y^2} & \text{sign}(z)\frac{(1-\alpha) * r * L}{g_y} \\ \text{sign}(z)\frac{\alpha * r * L}{g_x} & \text{sign}(z)\frac{(1-\alpha) * r * L}{g_y} & \frac{r}{|z|}(1+TL) \end{bmatrix}, \tag{18}$$

where

$$L = \frac{2(|z|-r)}{|z| - T(|z|-2r)}, \qquad T = \frac{\alpha x}{g_x} + \frac{(1-\alpha)y}{g_y},$$
$$g_x = \sqrt{x^2 + 4\alpha r(|z|-r)}, \qquad g_y = \sqrt{y^2 + 4(1-\alpha)r(|z|-r)}.$$

Finally, note that the power cone can be represented more generally as

$$\mathcal{K} = \{x \in \mathbf{R}^n \mid \prod_{i=1}^{m} x_i^{\alpha_i} \geq \|x_{m+1:n}\|, \ \ x_1, ..., x_m \geq 0\}, \tag{19}$$

where $\sum \alpha_i = 1$. However this can be modeled as a composition of three dimensional cones and a second-order cone, so we only concentrate on the three dimensional case.[1]

References

[1] Alnur Ali, Eric Wong, and J Zico Kolter. "A semismooth Newton method for fast, generic convex programming". In: *arXiv preprint arXiv:1705.00772* (2017).

[2] MOSEK ApS. *The MOSEK optimization toolbox for MATLAB manual. Version 9.0.* 2019. URL: `http://docs.mosek.com/9.0/toolbox/index.html`.

[3] Stephen Boyd and Lieven Vandenberghe. *Convex optimization*. Cambridge university press, 2004.

[4] Enzo Busseti, Walaa M Moursi, and Stephen Boyd. "Solution refinement at regular points of conic problems". In: *Computational Optimization and Applications* 74.3 (2019), pp. 627–643.

[5] Le Thi Khanh Hien. "Differential properties of Euclidean projection onto power cone". In: *Mathematical Methods of Operations Research* 82 (2015), pp. 265–284.

[6] Neal Parikh and Stephen Boyd. "Proximal algorithms". In: *Foundations and Trends in optimization* 1.3 (2014), pp. 127–239.

[7] Projection, presolve in MOSEK: exponential, and power cones. "Henrik A. Friberg". 23rd International Symposium on Mathematical Programming. 2018. URL: `https://docs.mosek.com/slides/2018/ismp2018/ismp-friberg.pdf`.

[1]See `https://docs.mosek.com/modeling-cookbook/powo.html#the-most-general-power-cone` for details.