Exponential cone projection

_proj_exp_cone() in diffcp explained

Axel Breuer

May 5, 2023

Abstract

Projection of a point $(\hat{x}, \hat{y}, \hat{z})$ onto the exponential cone $\{x, y, z | y \exp(x/y) \leq z, y \geq 0\}$ is a fundamental atomic operation used in conic optimization. The open-source solver `SCS` implemented such a projection via an ad-hoc iterative algorithm `proj_exp_cone` which was later incorporated and adapted into `diffcp` under the name `_proj_exp_cone`. The resulting code is compact and elegant, but, because I couldn't find a note explaining the algorithm, I ended up writing this short note to give an overview of its inner mechanics[1].

1 One problem, multiple formulations

For a given point $(\hat{x}, \hat{y}, \hat{z})$, finding its exponential cone projection consists in solving an equality constrained least squares problem[2]:

$$\begin{array}{ll} \min_{x,y,z} & \frac{1}{2}((x-\hat{x})^2 - (y-\hat{y})^2 - (z-\hat{z})^2) \\ \text{s.t.} & y\exp(x/y) = z \end{array}$$

The Lagrangian of this optimization problem is

$$\frac{1}{2}((x-\hat{x})^2 - ((y-\hat{y})^2 - (z-\hat{z})^2) + \mu(y\exp(x/y) - z) \quad (1)$$

where μ is the dual variable of the constraint $y\exp(x/y) = z$. The KKT condition at the solution (x^*, y^*, z^*, μ^*) are

$$x^* - \hat{x} + \mu^* \exp(x^*/y^*) = 0 \quad (2)$$

$$y^* - \hat{y} + \mu^* \exp(x^*/y^*)(1 - x^*/y^*) = 0 \quad (3)$$

$$z^* - \hat{z} - \mu^* = 0 \quad (4)$$

$$y^* \exp(x^*/y^*) - z^* = 0 \quad (5)$$

An equivalent formulation to the original minimization problem is

$$\begin{array}{ll} \min_{x,y,z} & \frac{1}{2}((x-\hat{x})^2 - (y-\hat{y})^2 - (z-\hat{z})^2) \\ \text{s.t.} & x + y\log(y/z) = 0 \end{array}$$

[1] I have tried to keep the names of the variables in the formulas as close as possible to the name of the variables in the C++ code. Known parameters are in blue .

[2] We have supposed that $(\hat{x}, \hat{y}, \hat{z})$ is strictly outside the exponential cone.

The Lagrangian of this optimization problem is

$$\frac{1}{2}((x-\hat{x})^2 - ((y-\hat{y})^2 - (z-\hat{z})^2) + \rho(x + y\log(y/z)) \tag{6}$$

where ρ is the dual variable of the constraint $x + y\log(y/z) = 0$. The KKT condition at the solution (x^*, y^*, z^*, μ^*) are

$$x^* - \hat{x} + \rho^* = 0 \tag{7}$$
$$y^* - \hat{y} + \rho^*(\log(y^*/z^*) + 1) = 0 \tag{8}$$
$$z^* - \hat{z} - \rho^* y^*/z^* = 0 \tag{9}$$
$$x^* + y^*\log(y^*/z^*) = 0 \tag{10}$$

2 Solving the ρ-formulation

We start with a few algebraic manipulations:

- Equation 7 implies

$$x^* = \hat{x} - \rho^* \tag{11}$$

- Equation 9 implies

$$y^* = \hat{z}(z^* - \hat{z})/\rho^* \tag{12}$$

- Replacing y^* in (8) by (12) gives

$$\frac{z^*(z^* - \hat{z})}{{\rho^*}^2} - \frac{\hat{y}}{\rho} + \log(\frac{z^* - \hat{z}}{\rho^*}) + 1 = 0 \tag{13}$$

which is equivalent to

$$\frac{(t^* + \hat{z})t^*}{{\rho^*}^2} - \frac{\hat{y}}{\rho} + \log(\frac{t^*}{\rho^*}) + 1 = 0 \tag{14}$$

where $t^* := z^* - \hat{z}$.

For a given value of ρ^*, the system of equations 11, 12 and 13 can be solved in the following order

1. Find t^* by solving (14) via a 1-dimensional Newton method.
2. Set $z^* := t^* + \hat{z}$ (which solves equation 13).
3. Set $y^* = \hat{z}(z^* - \hat{z})/\rho^*$ (which is equation 12).
4. Set $x^* = \hat{x} - \rho^*$ (which is equation 11).

The above equations explicitly show us how the 3 first equations of the KKT system (7, 8, 9) express (x^*, y^*, z^*) as a function of ρ^*.

In order to identify all four parameters (x^*, y^*, z^*, ρ^*), we have now to use the fourth equation of the KKT system (10). This is performed via a bisection algorithm, which returns ρ^* such that $x^*(\rho^*) + y^*(\rho^*)\log(y^*(\rho^*)/z^*(\rho^*)) = 0$