PROJECTION ONTO A SIMPLEX*

Abstract. Explains how to find a projection onto a simplex.

1. Projection onto a face($n-1$ simplex). Consider a projection problem as below:

$$\min_{x} \quad \frac{1}{2}\|x-z\|_2^2 \tag{1.1a}$$

$$\text{s.t.} \quad e^T x \;=\; b \tag{1.1b}$$

$$x \;\geq\; 0. \tag{1.1c}$$

where $e \in \mathbb{R}^n$ is a vector of all 1's. Then the KKT condition of (1.1a) is

$$0 \;\leq\; x - z + \lambda e \;\perp\; x \;\geq\; 0 \tag{1.2}$$

where λ is the Largrange multiplier to (1.1b). So the solution x^* of (1.1a) is

$$x^* \;=\; (z-\lambda e)_+ \;=\; \max\{z-\lambda e,\; 0\} \tag{1.3}$$

where the max operator applies element-wise. If we define a function $g(\lambda)$ like below

$$g(t) \;:=\; \sum_{i:z_i-\lambda\geq 0}(z_i-\lambda) \;=\; \sum_{i=1}^{n} x_i \;=\; e^T x \tag{1.4}$$

then the optimal solution x^* of (1.1a) can be obtained by finding a value of λ^* such that $g(\lambda^*)=b$.

Without loss of generality, we assume that the vector z is sorted in descending order. Let w be a vector such that

$$w_k \;=\; \sum_{i=1}^{k} z_i. \tag{1.5}$$

Assume that λ_k is the solution of $g(\lambda)=b$ when the first k entries of z have $z_i - \lambda^k \geq 0$. Then from

$$\sum_{i=1}^{k}(z_i-\lambda_k) \;=\; b \tag{1.6}$$

we have

$$\lambda_k \;=\; \frac{w_k-b}{k}. \tag{1.7}$$

And we need to find the k^* such that $z_{k^*}-\lambda_{k^*} \geq 0$ and $z_{k^*+1}-\lambda_k^* \leq 0$. Then $\lambda^* = \lambda_{k^*}$ and $x^* = (z-\lambda^* e)_+$.

*June 3, 2013

2. Projection onto n-simplex. Now we consider a projection onto n-simplex:

$$\min_x \quad \frac{1}{2}\|x - z\|_2^2 \tag{2.1a}$$

$$\text{s.t.} \quad e^T x \ \le \ b \tag{2.1b}$$

$$x \ \ge \ 0. \tag{2.1c}$$

The KKT condition of (2.1a) is

$$0 \ \le \ x - z + \lambda e \ \perp \ x \ \ge \ 0 \tag{2.2a}$$

$$0 \ \le \ b - e^T x \ \perp \ \lambda \ \ge \ 0. \tag{2.2b}$$

Again, λ is the Lagrange multiplier.

First, if $e^T z_+ \le b$, then it can be easily shown that $x^* = z_+$ is a solution. Thus assume that $e^T z_+ > b$. Then the solution should be $x^* = (z - \lambda^* e)_+$ from the KKT condition where λ^* is the optimal Lagrange multiplier. Also we can see that $e^T x^* = b$. If not, i.e. $e^T x^* < b$, we should have $\lambda^* = 0$ and

$$b \ > \ e^T x^* \ = \ e^T(z - \lambda^* e)_+ \ = \ e^T z_+ \ > \ b \tag{2.3}$$

which contradicts. Thus we can use the same technique discussed in Section 1. Note that since $e^T z_+ > b$, λ^* will be nonnegative if the smallest k^* is chosen, and thus satisfies the KKT condition.

3. Projection onto an intersection of n-simplex and box constraints. We consider another problem like below:

$$\min_x \quad \frac{1}{2}\|x - z\|_2^2 \tag{3.1a}$$

$$\text{s.t.} \quad e^T x \ \le \ b \tag{3.1b}$$

$$0 \ \le \ x \ \le \ u. \tag{3.1c}$$

Since the constraints are convex, we can solve this problem using Dykstra's projection algorithm. Let $C = \{x \mid e^T x \le b\}$ and $D = \{x \mid 0 \le x \le u\}$. Then the Dykstra's projection algorithm is described in algorithm 1.[1] To use the algorithm, we need to know the projection onto each convex set C and D. For (3.1b), we can use the projection described in previous sections. Since (3.1c) is a simplex box constraint, we can simply truncate to project onto this set.

[1] The algorithm is from the wiki page. `http://en.wikipedia.org/wiki/Dykstra's_projection_algorithm`

Algorithm 1 PROJECTION ONTO $C \cap D$

Require: z: A point projected.
C, D: Convex Sets.
$\mathcal{P}_C(x), \mathcal{P}_D(x)$: Projection functions onto C and D, respectively.
Ensure: x: Projection of z onto $C \cap D$.

$k \leftarrow 0$
$x_0 \leftarrow z$, $y_0 \leftarrow 0$, $p_0 \leftarrow 0$, $q_0 \leftarrow 0$

repeat
  $y_k \leftarrow \mathcal{P}_D(x_k + p_k)$
  $p_{k+1} \leftarrow (x_k - y_k) + p_k$
  $x_{k+1} \leftarrow \mathcal{P}_C(y_k + q_k)$
  $q_{k+1} \leftarrow (y_k - x_{k+1}) + q_k$
  $k \leftarrow k + 1$
until x_k, y_k, p_k and q_k are fixed points.