Homework 4

Ilia Kamyshev

March 5, 2021

Contents

1 Euclidean Ball **1**

2 Projection to the hyperplane **2**

3 Conjugate function **3**

4 Primal problem **3**
- 4.1 Derivation . . . 3
- 4.2 Solution . . . 4

5 Optimization problem **4**
- 5.1 Analysis of primal problem . . . 4
- 5.2 Lagrangian and dual function: . . . 4
- 5.3 The dual problem . . . 6

6 Quadratic programming problem **6**
- 6.1 The dual function . . . 6
- 6.2 Primal from dual . . . 7
- 6.3 ADMM implementation . . . 7
- 6.4 Decomposition . . . 7

1 Euclidean Ball

Given

$$\min_{x\in\mathbb{R}^n:||x-c||\leq r} ||z-x||$$

Making power of 2 of the objective function and the constraints will not affect the solution, just for simplicity of taking gradients. Thus, the Lagrangian:

$$\mathcal{L}(x,\lambda) = ||x-z||^2 + \lambda(||x-c||^2 - r^2),\ \lambda \text{ is scalar since } ||x-c||^2 - r^2 \text{ is also scalar}$$

The dual function $g(\lambda) = \min_x \mathcal{L}(x,\lambda)$. Thus, the gradient:

$$\nabla_x \mathcal{L}(x,\lambda)|_{x=x^*} = 2(x^*-z) + \lambda(2(x^*-c)) = 0 \implies x^* = \frac{\lambda c + z}{\lambda + 1}$$

The dual problem:

$$d^* = \max_\lambda g(\lambda) = \max_\lambda \{||\frac{\lambda c + z}{\lambda+1} - z||^2 + \lambda(||\frac{\lambda c+z}{\lambda+1} - c||^2 - r^2)\}$$

Or

$$d^* = \max_\lambda\{\frac{\lambda^2}{(\lambda+1)^2}||c-z||^2 + \lambda(\frac{1}{(\lambda+1)^2}||c-z||^2 - r^2)\} = \max_\lambda\{\frac{1}{(\lambda+1)^2}||c-z||^2(\lambda^2+\lambda) - \lambda r^2\} =$$

$$= \max_\lambda\{\frac{\lambda}{\lambda+1}||c-z||^2 - \lambda r^2\}$$

Therefore,

$$\nabla_\lambda g(\lambda)|_{\lambda=\lambda^*} = \frac{\lambda^*+1-\lambda^*}{(\lambda^*+1)^2}||c-z||^2 - r^2 = 0 \implies (\lambda^*+1)^2 = \frac{||c-z||^2}{r^2} \implies \lambda^* = \frac{||c-z||}{r} - 1$$

Here we took only positive root since $\lambda \geq 0$. The dual problem d^*:

$$\frac{\frac{||c-z||}{r} - 1}{\frac{||c-z||}{r}}||c-z||^2 - (\frac{||c-z||}{r} - 1)r^2 = ||c-z||^2 - 2r||c-z|| + r^2 = (||c-z|| - r)^2$$

Answer: $d^* = (||c-z||-r)^2$, $x^* = \frac{\lambda^* c + z}{\lambda^*+1}$, $\lambda^* = \frac{||c-z||}{r} - 1$

2 Projection to the hyperplane

Given

$$\min_{x\in q:(a,x)=b} ||x-z||$$

The Lagrangian:

$$\mathcal{L}(x,v) = ||x-z||^2 + v((a,x)-b),\ v \text{ is scalar since } (a,x)=b \text{ is scalar too}$$

$$\nabla_x \mathcal{L}(x,v)|_{x=x^*} = 2(x^*-z) + va = 0$$

$$x^* = z - v\frac{a}{2}$$

The dual function is:

$$g(v) = \min_x \mathcal{L}(x,v) = ||z - v\frac{a}{2} - z||^2 + v((a, z - v\frac{a}{2}) - b)$$

$$g(v) = (v\frac{a}{2}, v\frac{a}{2}) + v((a, z - v\frac{a}{2}) - b) = (v\frac{a}{2}, v\frac{a}{2}) + v(a^t z - a^t a\frac{v}{2} - b) = \frac{v^2}{4}(a,a) + v((a,z) - \frac{v}{2}(a,a) - b)$$

$$g(v) = \frac{v^2}{2}(a,a)(\frac{1}{2} - 1) - vb + v(a,z) = -\frac{v^2}{4}(a,a) + v((a,z) - b)$$

The dual problem is:

$$d^* = \max_v g(v)$$

$$\nabla_v g(v)|_{v=v^*} = -\frac{v^*}{2}(a,a) + (a,z) - b = 0 \implies v^* = 2\frac{(a,z)-b}{(a,a)}$$

Thus,

$$d^* = g(v^*) = -4\frac{((a,z)-b)^2}{4(a,a)^2}(a,a) + 2\frac{(a,z)-b}{(a,a)}((a,z)-b)$$

$$d^* = -\frac{((a,z)-b)^2}{(a,a)} + 2\frac{((a,z)-b)^2}{(a,a)}$$

$$d^* = \frac{((a,z)-b)^2}{(a,a)}$$

Answer: $d^* = \frac{((a,z)-b)^2}{(a,a)}$, $x^* = z - v^*\frac{a}{2}$, $v^* = 2\frac{(a,z)-b}{(a,a)}$

3 Conjugate function

Given

$$\min_{Ax=b} f(x)$$

And conjugate function:

$$f^*(y) = \sup_{x\in\mathbb{R}^n} ((x,y) - f(x))$$

So the Lagrangian is given by:

$$\mathcal{L}(x,v) = f(x) + v^T(Ax-b),\ v \text{ is a vector, since } Ax-b \text{ also vector}$$

The dual function:

$$g(v) = \inf_x \mathcal{L}(x,v) = \inf_x\{f(x)+(Ax-b)^T v\} = \inf_x \mathcal{L}(x,v) = \inf_x\{f(x)+x^T A^T v - b^T v\} = \inf_x\{f(x)+(x,A^T v)-(b,v)\}$$

Taking into account our conjugate function in terms of $y = A^T v$:

$$g(v) = \inf_x\{f(x) + (A^T v, x) - (b,v)\} = -\inf_x\{-f(x) + (-A^T v, x) + (b,v)\} = -f^*(-A^T v) - (b,v)$$

The domain:

$$\textbf{dom } g(v) = \{v : -A^T v \in \textbf{dom } f^*\}$$

Answer: a) $-f^*(-A^T v) - (b,v)$ b) $\textbf{dom } g(v) = \{v : -A^T v \in \textbf{dom } f^*\}$

4 Primal problem

Given

$$\min_{(ax,x)\leq 1} (c,x)$$

4.1 Derivation

The Lagrangian:

$$\mathcal{L}(x,\lambda) = (c,x) + \lambda((ax,x)-1),\ \lambda \text{ is scalar since } (ax,x) \text{ is scalar}$$

The dual function:

$$g(\lambda) = \min_x \mathcal{L}(x,\lambda)$$

Thus, the gradient:

$$\nabla_x \mathcal{L}(x,\lambda)|_{x=x^*} = c + 2\lambda a x^* = 0 \implies x^* = -\frac{1}{2\lambda}a^{-1}c$$

Finally, the dual function:

$$g(\lambda) = -\frac{1}{2\lambda}(c, a^{-1}c) + \lambda(\frac{1}{4\lambda^2}(aa^{-1}c, a^{-1}c) - 1)$$

$$g(\lambda) = \frac{1}{2\lambda}(c, a^{-1}c)(\frac{1}{2} - 1) - \lambda = -\frac{1}{4\lambda}(c, a^{-1}c) - \lambda$$

And the dual problem:

$$d^* = \max_\lambda g(\lambda) = \max_\lambda\{-\frac{1}{4\lambda}(c, a^{-1}c) - \lambda\}$$

4.2 Solution

The gradient over λ:

$$\nabla_\lambda g(\lambda)|_{\lambda=\lambda^*} = \frac{1}{4\lambda^{*2}}(c, a^{-1}c) - 1 = 0 \implies \lambda^* = \frac{1}{2}\sqrt{(c, a^{-1}c)} \text{ since } \lambda \geq 0$$

Thus, the primal problem has a solution through the dual problem:

$$d^* = -\frac{1}{2\sqrt{(c, a^{-1}c)}}(c, a^{-1}c) - \frac{1}{2}\sqrt{(c, a^{-1}c)} = -\sqrt{(c, a^{-1}c)}$$

Answer: $d^* = -\sqrt{(c, a^{-1}c)}$

5 Optimization problem

Given

$$\min_x \{x^2 + 1\}$$

Subject to $(x-2)(x-4) \leq 0$.

5.1 Analysis of primal problem

The feasible set is $x \in [2, 4]$ (since $(x-2)(x-4) \leq 0$). Optimal solution x^* can be obtained from $(x^2+1)' = 0$ as $x^* = 0$, thus the optimal value $p^* = 1$ – contradiction with constraints. The very left available $x = 2$, indeed, from the plot 1 we can see that this argument corresponds to the lowest value of the function within given feasible region, thus $p^* = 2^2 + 1 = 5$.

5.2 Lagrangian and dual function:

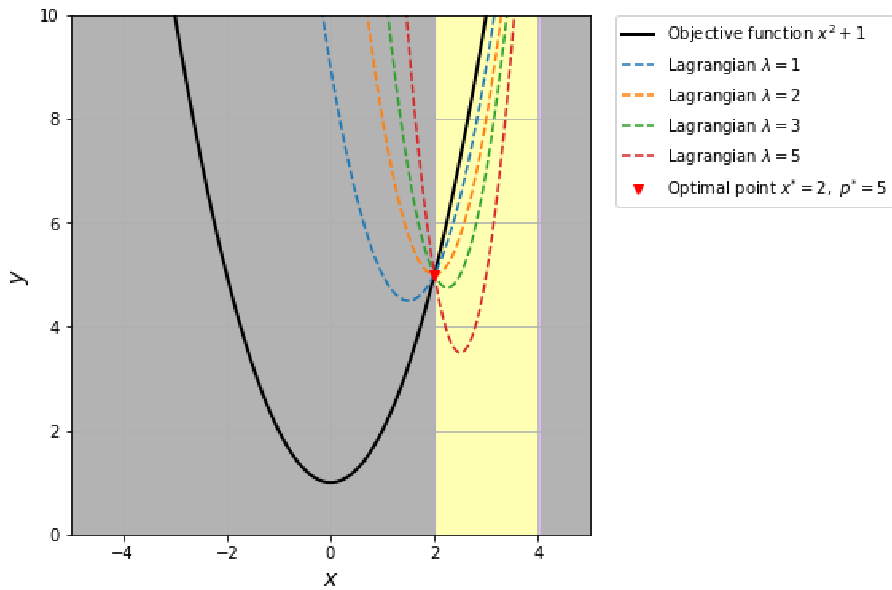

Figure 1: Lagrangians for different λ and the objective function. Feasible region is filled with yellow color.

The Lagrangian:

$$\mathcal{L}(x, \lambda) = x^2 + 1 + \lambda(x-2)(x-4) \text{ with scalar } \lambda$$

The dual function:

$$g(\lambda) = \min_x \mathcal{L}(x, \lambda)$$

The optimal solution:

$$\nabla_x \mathcal{L}(x, \lambda)|_{x=x^*} = 2x^* + \lambda(x^* - 4) + \lambda(x^* - 2) = 0 \implies x^* = \frac{3\lambda}{1+\lambda}$$

Now the dual function is:

$$g(\lambda) = (\frac{3\lambda}{1+\lambda})^2 + 1 + \lambda(\frac{3\lambda}{1+\lambda} - 2)(\frac{3\lambda}{1+\lambda} - 4) =$$

$$(\frac{3\lambda}{1+\lambda})^2 + 1 - \lambda(\frac{\lambda-2}{1+\lambda})(\frac{\lambda+4}{1+\lambda}) =$$

$$= \frac{1}{(1+\lambda)^2}(9\lambda^2 + 1 + 2\lambda + \lambda^2 - \lambda(\lambda^2 + 2\lambda - 8)) =$$

$$= \frac{1}{(1+\lambda)^2}(8\lambda^2 + 10\lambda - \lambda^3 + 1) = \frac{1}{(1+\lambda)^2}(\lambda+1)(-\lambda^2 + 9\lambda + 1) =$$

$$= \frac{1}{(1+\lambda)}(-\lambda^2 + 9\lambda + 1 + 8\lambda^2 - 8\lambda^2) = \frac{1}{(1+\lambda)}(-9\lambda^2 + 9\lambda + 1 + 8\lambda^2) =$$

$$= \frac{1}{(1+\lambda)}(-9\lambda^2 + (\lambda+1)(8\lambda+1)) = \frac{-9\lambda^2}{\lambda+1} + 8\lambda + 1$$

Lower bound property holds since all the peaks of Lagrangians (for $\lambda \geq 0$) are below p^* (see plot 2). In the next paragraph you will also see that from optimal solution λ^* the values of $g(\lambda)$ can't be more than $p^* = 5$. Even for $\lambda \to \infty$ the limit $\lim_{\lambda \to \infty} g(\lambda) = -\infty$.

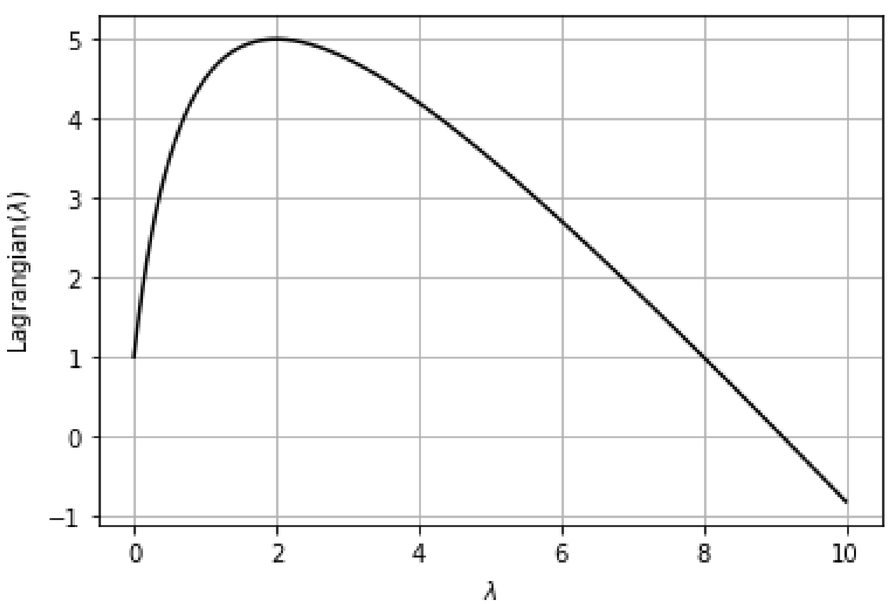

Figure 2: Dual function for $\lambda \geq 0$

5.3 The dual problem

states that:

$$d^* = \max_{\lambda} g(\lambda)$$

Now we can compute the gradient over λ:

$$\nabla_\lambda g(\lambda) = \frac{-18\lambda(\lambda+1) + 9\lambda^2}{(\lambda+1)^2} + 8 = \frac{-9\lambda^2 - 18\lambda}{(\lambda+1)^2} + 8 = -9\lambda\frac{\lambda+2}{(\lambda+1)^2} + 8 = 0$$

$$9\lambda(\lambda+2) = 8(\lambda^2 + 2\lambda + 1)$$

$$\lambda^2 + 2\lambda - 8 = 0 \implies \lambda^* = 2 \text{ since } \lambda \geq 0, \text{ the other is } -4 \text{ which is omitted}$$

Thus,

$$d^* = \frac{-36}{3} + 16 + 1 = 5$$

Strong duality holds if $p^* = d^*$, in our case $p^* = 5 = d^*$ this holds.

6 Quadratic programming problem

Given

$$\min_x \frac{1}{2} x^T P x + (x, q)$$

Subject to $Ax \leq b$.

6.1 The dual function

$$g(\lambda) = \min_x \frac{1}{2} x^T P x + (x, q) + \lambda^T (Ax - b) \ \lambda \text{ is vector since } Ax - b \text{ is vector too}$$

The Lagrangian can be rewritten as:

$$\mathcal{L}(x, \lambda) = \frac{1}{2} x^T P x + x^T q + x^T A^T \lambda - \lambda^T b$$

The gradient over x:

$$\nabla_x \mathcal{L}(x, \lambda)|_{x=x^*} = Px^* + q + A^T\lambda = 0 \implies x^* = -P^{-1}q - P^{-1}A^T\lambda$$

Therefore,

$$g(\lambda) = \frac{1}{2}(-P^{-1}q - P^{-1}A^T\lambda)^T P(-P^{-1}q - P^{-1}A^T\lambda) + (-P^{-1}q - P^{-1}A^T\lambda)^T q + (-P^{-1}q - P^{-1}A^T\lambda)^T A^T\lambda - \lambda^T b =$$

$$= (-P^{-1}q - P^{-1}A^T\lambda)^T(\frac{1}{2}P(-P^{-1}q - P^{-1}A^T\lambda) + q + A^T\lambda) - \lambda^T b = (-P^{-1}q - P^{-1}A^T\lambda)^T(-\frac{1}{2}q - \frac{1}{2}A^T\lambda + q + A^T\lambda) - \lambda^T b$$

$$= (-P^{-1}q - P^{-1}A^T\lambda)^T(\frac{1}{2}q + \frac{1}{2}A^T\lambda) - \lambda^T b = -\frac{1}{2}(P^{-1}(q + A^T\lambda))^T(q + A^T\lambda) - \lambda^T b =$$

$$= -\frac{1}{2}(q + A^T\lambda)^T P^{-1}(q + A^T\lambda) - \lambda^T b$$

The optimal λ^*:

$$\nabla_\lambda g(\lambda) = -\frac{1}{2}[\frac{\partial}{\partial\lambda}(q + A^T\lambda)](2P^{-1}(q + A^T\lambda)) - b = -AP^{-1}(q + A^T\lambda) - b = 0$$

$$-AP^{-1}A^T\lambda = b + AP^{-1}q$$

$$\lambda^* = -(AP^{-1}A^T)^{-1}(b + A^T P^{-1} q)$$

The dual problem:

$$d^* = g(\lambda^*)$$

6.2 Primal from dual

Primal optimal solution via λ^* is $x^* = -P^{-1}q + P^{-1}A^T(AP^{-1}A^T)^{-1}(b + A^TP^{-1}q) = A^{-1}b$

6.3 ADMM implementation

At first, we should have augmented Lagrangian:

$$\mathcal{L}(\lambda, z, y) = -\lambda^T b - \frac{1}{2}(q + A^T\lambda)^T P^{-1}(q + A^T\lambda) + g(z) + y^T(\lambda - z) + \frac{1}{2}\rho||\lambda - z||^2$$

Let's assign $Q = \{\lambda : \lambda \geq 0\}$. Keep in mind, that ADMM solves $\min_{\lambda=z} f(\lambda) + I_Q(z)$, where $I_Q(z) = 0$ if $z \in Q$ and $I_Q(z) = \infty$ otherwise. Then, the ADMM iterations are:

$$\lambda(k+1) = \text{argmin}_\lambda(-\lambda^T b - \frac{1}{2}(q + A^T\lambda)^T P^{-1}(q + A^T\lambda) + \frac{1}{2}\rho||\lambda - z(k) + u(k)||^2)$$

$$z(k+1) = \text{argmin}_z(g(z) + \frac{1}{2}\rho||\lambda(k+1) - z + u(k)||^2) = \mathbf{Proj}_Q(\lambda(k+1) + u(k))$$

$$u(k+1) = u(k) + \lambda(k+1) - z(k+1)$$

Obviously, that $u(k+1) = u(k) + \lambda(k+1) - \mathbf{Proj}_Q(\lambda(k+1) + u(k)) = u(k) + \lambda(k+1) - \max(\lambda(k+1) + u(k), 0) = -\max(-u(k) - \lambda(k+1), 0)$.

6.4 Decomposition

$$g(\lambda) = -\frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n} P_{ij}^{-1}(q + A^T\lambda)_i(q + A^T\lambda)_j - \lambda^T b$$

As in [Boyd] $\lambda^t b$ is an indicator of convex set. So:

$$g(\alpha, \beta) = -\frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n} P_{ij}^{-1}\alpha_i\alpha_j - \beta^T b$$

Where the first term is separable since it is a sum of functions of individual variables $\alpha_{i,j}$.