236330 - Introduction to Optimization: Homework #4

June 24, 2020

Amit Rotner
123456789
Or Steiner
123456789

DNNs, Augmented Lagrangian and Constrained Optimization

Part 2 – Augmented Lagrangian:

Task 1: Consider the following quadratic problem:

$$\min_{x} 2\left(x_1-5\right)^2+\left(x_2-1\right)^2$$

s.t.

$$\begin{aligned} x_2 \leq & 1-\frac{x_1}{2} \\ x_2 \geq & x_1 \\ x_2 \geq & -x_1 \end{aligned}$$

1. Is this problem convex?
2. Find the optimal solution (x^*, λ^*)
3. Calculate the Lagrange multipliers using KKT conditions. What is the optimal value of the objective function?

Solution:

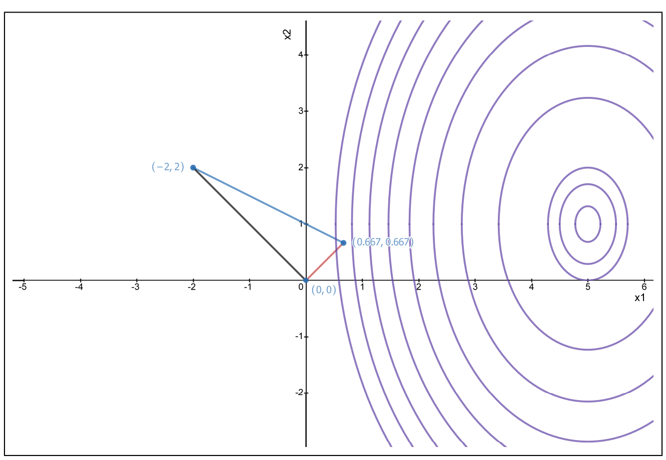


Figure 1: The feasible area and the level sets of the objective function

1. Therefore the active constraints are c_1 and c_2.

2. It is fairly clear from the figure above that the solution is $x^* = \left(\frac{2}{3}, \frac{2}{3}\right)^T$.

3. Writing the Lagrangian yields:

$$\mathcal{L}(x, \lambda) = 2(x_1 - 5)^2 + (x_2 - 1)^2 + \lambda_1 \left(x_2 - 1 + \frac{x_1}{2}\right) + \lambda_2 (x_1 - x_2) + \lambda_3 (-x_1 - x_2)$$

First we find the gradient of the Lagrangian:

$$\nabla_x \mathcal{L}(x, \lambda) = \left(4(x_1 - 5) + 0.5\lambda_1 + \lambda_2 - \lambda_3, \quad 2(x_2 - 1) + \lambda_1 - \lambda_2 - \lambda_3\right)$$

If x^* is an optimal solution for our problem, then there exist λ^* which satisfies the KKT conditions:

(a) $\nabla_{x_1} \mathcal{L}(x^*, \lambda) = 0 \Rightarrow 4(x_1 - 5) + 0.5\lambda_1 + \lambda_2 - \lambda_3 = 0$
(b) $\nabla_{x_2} \mathcal{L}(x^*, \lambda) = 0 \Rightarrow 2(x_2 - 1) + \lambda_1 - \lambda_2 - \lambda_3 = 0$
(c) $x_2 - 1 + \frac{x_1}{2} \leq 0$
(d) $x_1 - x_2 \leq 0 \Rightarrow x_1 \leq x_2$
(e) $-x_1 - x_2 \leq 0 \Rightarrow -x_2 \leq x_1$
(f) $\lambda^* \geq 0$
(g) $\lambda_1 \left(x_2 - 1 + \frac{x_1}{2}\right) = 0$
(h) $\lambda_2 (x_1 - x_2) = 0$
(i) $\lambda_3 (-x_1 - x_2) = 0$

From (9) we get:

$$\lambda_3 = 0$$

From $(1) + (2)$ we get:

$$0.5\lambda_1 + \lambda_2 = 17\frac{1}{3}$$
$$\lambda_1 - \lambda_2 = \frac{2}{3}$$

Therefore, we get $\lambda_1 = 12, \lambda_2 = 11\frac{1}{3}$.

In conclusion:

$$x^* = \left(\frac{2}{3}, \frac{2}{3}\right)^T, \ \lambda^* = \left(12, 11\frac{1}{3}, 0\right)^T$$

Hence, the optimal value of the objective function is:

$$2\left(\frac{2}{3} - 5\right)^2 + \left(\frac{2}{3} - 1\right)^2 = 37\frac{2}{3}$$

4. From the above, we know that:

$$\mathcal{L}(x,\lambda) = 2(x_1-5)^2 + (x_2-1)^2 + \lambda_1\left(x_2 - 1 + \frac{x_1}{2}\right) + \lambda_2(x_1-x_2) + \lambda_3(-x_1-x_2)$$

If we hold $\lambda_1, \lambda_2, \lambda_3$ fixed, this is a convex function of $(x_1,x_2)^T$.
Therefore, the infimum with respect to $(x_1,x_2)^T$ is achieved when the partial derivatives with respect to x_1 and x_2 are zero, that is,

$$4(x_1-5) + 0.5\lambda_1 + \lambda_2 - \lambda_3 = 0$$

$$2(x_2-1) + \lambda_1 - \lambda_2 - \lambda_3 = 0$$

Hence,

$$x_1 = \frac{20 - 0.5\lambda_1 - \lambda_2 + \lambda_3}{4}$$

$$x_2 = \frac{2 - \lambda_1 + \lambda_2 + \lambda_3}{2}$$

By substituting these infimal values into $\mathcal{L}(x,\lambda)$ we obtain the dual objective:

$$\begin{aligned}\mathcal{L}(x,\lambda) &= 2(x_1-5)^2 + (x_2-1)^2 + \lambda_1\left(x_2 - 1 + \frac{x_1}{2}\right) + \lambda_2(x_1-x_2) + \lambda_3(-x_1-x_2) \\ &= 2\left(\frac{20 - 0.5\lambda_1 - \lambda_2 + \lambda_3}{4} - 5\right)^2 + \left(\frac{2 - \lambda_1 + \lambda_2 + \lambda_3}{2} - 1\right)^2 \\ &+ \lambda_1\left(\frac{2 - \lambda_1 + \lambda_2 + \lambda_3}{2} - 1 + \frac{20 - 0.5\lambda_1 - \lambda_2 + \lambda_3}{8}\right) + \lambda_2\left(\frac{20 - 0.5\lambda_1 - \lambda_2 + \lambda_3}{4} - \frac{2 - \lambda_1 + \lambda_2 + \lambda_3}{2}\right) \\ &+ \lambda_3\left(-\frac{20 - 0.5\lambda_1 - \lambda_2 + \lambda_3}{4} - \frac{2 - \lambda_1 + \lambda_2 + \lambda_3}{2}\right)\end{aligned}$$

5. Hence, the dual problem is:

$$\begin{aligned}\underset{\lambda\geq 0}{max} 2&\left(\frac{20 - 0.5\lambda_1 - \lambda_2 + \lambda_3}{4} - 5\right)^2 + \left(\frac{2 - \lambda_1 + \lambda_2 + \lambda_3}{2} - 1\right)^2 \\ &+ \lambda_1\left(\frac{2 - \lambda_1 + \lambda_2 + \lambda_3}{2} - 1 + \frac{20 - 0.5\lambda_1 - \lambda_2 + \lambda_3}{8}\right) + \lambda_2\left(\frac{20 - 0.5\lambda_1 - \lambda_2 + \lambda_3}{4} - \frac{2 - \lambda_1 + \lambda_2 + \lambda_3}{2}\right) \\ &+ \lambda_3\left(-\frac{20 - 0.5\lambda_1 - \lambda_2 + \lambda_3}{4} - \frac{2 - \lambda_1 + \lambda_2 + \lambda_3}{2}\right)\end{aligned}$$

6. Substitute the Lagrange multipliers from section 3 into the dual problem, yields that the value of the dual problem is $= 37\frac{2}{3}$ which it also the optimum of the dual problem.
Since it is also the optimal value of the primal problem, the duality gap is zero.

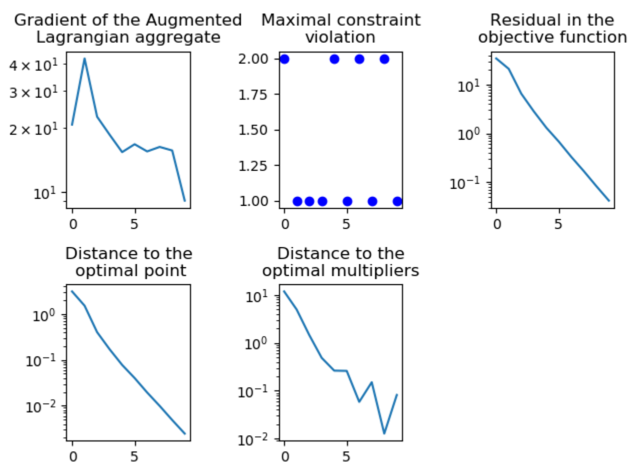
Task 2: Writing the quadratic programming problem used in Task 1 in general form, yields:

$$\frac{1}{2}\begin{pmatrix}x_1 & x_2\end{pmatrix}\underbrace{\begin{pmatrix}4 & 0\\ 0 & 2\end{pmatrix}}_{Q}\begin{pmatrix}x_1\\ x_2\end{pmatrix}+\underbrace{\begin{pmatrix}-20 & -2\end{pmatrix}}_{d^T}\begin{pmatrix}x_1\\ x_2\end{pmatrix}+\underbrace{51}_{e}$$

s.t.

$$\underbrace{\begin{pmatrix}0.5 & 1\\ 1 & -1\\ -1 & -1\end{pmatrix}}_{A}\begin{pmatrix}x_1\\ x_2\end{pmatrix}-\underbrace{\begin{pmatrix}1\\ 0\\ 0\end{pmatrix}}_{b}\leq 0$$

Using the Augmented Lagrangian Solver to Solve a Quadratic Programming Problem:



Part 3 – Constrained Optimization:

Task 1: Consider the following optimization problem:

$$\min_{x}\ x^T M x + c^T x$$

s.t.

$$Ax = b$$
$$M \succ 0$$
$$A \in \mathbb{R}^{m\times n}, x \in \mathbb{R}^n, c \in \mathbb{R}^n, b \in \mathbb{R}^n, M \in \mathbb{R}^{m\times m}$$

1. Is this problem convex?
2. Find the optimal solution (x^*, λ^*)

Solution:

1. We start by finding the Hessian of the objective function:

$$\begin{aligned} df &= dx^T Mx + x^T M dx + c^T dx \\ &= x^T M^T dx + x^T M dx + c^T dx \\ &\underset{(i)}{=} 2x^T M dx + c^T dx \\ &\underset{(ii)}{=} \left(2x^T M + c^T\right) dx \\ &= (2Mx + c)^T dx \end{aligned}$$

   (i) - Since $dx^T Mx$ is a scalar.

   (ii) - Since $M = M^T$.

   And therefore: $\nabla f = 2Mx + c$

   Now,

$$d\nabla f = 2M dx$$

   Hence, $H = 2M$.

   We know that $M \succ 0$ and therefore $H = 2M$ is also PD and the problem is convex.

2. Writing the Lagrangian yields:

$$\mathcal{L}(x, \lambda) = f(x) + \lambda (Ax - b)$$

   First we find the gradient of the Lagrangian:

$$\nabla_x \mathcal{L}(x, \lambda) = 2Mx + c + \lambda A^T$$

   If x^* is an optimal solution for our problem, then there exist λ^* which satisfies the KKT conditions:

   (a) $\nabla_x \mathcal{L}(x, \lambda) = 0 \Rightarrow 2Mx^* + \lambda A^T = -c$
   (b) $Ax^* - b = 0$

   Writing the above as a linear system:

$$\begin{pmatrix} 2M & A^T \\ A & 0 \end{pmatrix} \begin{pmatrix} x^* \\ \lambda^* \end{pmatrix} = \begin{pmatrix} -c \\ b \end{pmatrix}$$

   Assuming A is invertible and since M is PD, then $\begin{pmatrix} 2M & A^T \\ A & 0 \end{pmatrix}$ is also invertible and:

$$\begin{pmatrix} x^* \\ \lambda^* \end{pmatrix} = \begin{pmatrix} 2M & A^T \\ A & 0 \end{pmatrix}^{-1} \begin{pmatrix} -c \\ b \end{pmatrix}$$

Task 2: Consider the following optimization problem:

$$\underset{x}{min}\ \|x-c\|_2^2$$

s.t.

$$Ax = b$$
$$A \in \mathbb{R}^{m\times n}, x \in \mathbb{R}^n, c \in \mathbb{R}^n, b \in \mathbb{R}^n$$

1. Is this problem convex?
2. Find the optimal solution (x^*, λ^*)

Solution:

1. We start by proving that the objective function is convex:
   We have seen in class that the Euclidean norm $\|\cdot\|_2$ is convex and therefore $h(x) = \|x-c\|_2$ is also convex.
   Since $g(x) = x^2$ is convex as well, and also it is non-decreasing on $[0, \infty)$, the range of h,
   the composition $f = f \circ h = \|x-c\|_2^2$ is convex.
   Hence, the problem is convex.
2. Writing the Lagrangian yields:
   $$\mathcal{L}(x,\lambda) = f(x) + \lambda(Ax-b)$$
   First we find the gradient of the Lagrangian:
   $$\nabla_x \mathcal{L}(x,\lambda) = 2x^T x - 2x^T c + \lambda A^T$$
   If x^* is an optimal solution for our problem, then there exist λ^* which satisfies the KKT conditions:
   (a) $\nabla_x \mathcal{L}(x,\lambda) = 0 \Rightarrow 2x^{*T}x^* - 2x^{*T}c = \lambda A^T$
   (b) $Ax^* - b = 0$
   Assuming A is invertible, then $x^* = A^{-1}b$.
   Hence,
   $$2b^T\left(A^{-1}\right)^T A^{-1}b - 2b^T\left(A^{-1}\right)^T c = \lambda A^T$$
   $$\lambda^* = 2b^T\left(AA^T\right)^{-1} b\left(A^T\right)^{-1} - 2b^T\left(A^{-1}\right)^T c\left(A^T\right)^{-1}$$
   Therefore, the optimal solution is:
   $$(x^*, \lambda^*) = \left(A^{-1}b, 2b^T\left(AA^T\right)^{-1} b\left(A^T\right)^{-1} - 2b^T\left(A^{-1}\right)^T c\left(A^T\right)^{-1}\right)$$