APM 523
NLP

Matt Kinsinger, Hongjun Choi, Adarsh Akkshai

November 1, 2017

1 KKT Conditions

The **Karush-Kuhn-Tucker(KKT) conditions** are first-order necessary conditions for a solution to be optimal and provided that some regularity conditions are satisfied. Especially, the KKT conditions can be used to solve inequality constraints that the KKT approach to nonlinear programming generalizes the method of Lagrange multipliers, which allows only equality constraints. In general, many optimization algorithm can be interpreted as methods for numerically solving the KKT system of equations are inequalities.

In our problem, we can consider following nonlinear minimization as below :

$$
\begin{aligned}
\min_x F = x_1^2 - 3x_2 - 4y_1 + y_2^2 \\
\text{subject to} \quad x_1^2 + 2x_2 \leq 4, \\
x_1 \geq 0, x_2 \geq 0, \\
\min_y 2x_1^2 + y_1^2 - 5y_2 \\
\text{subject to} \quad x_1^2 - 2x_1 + x_2^2 - 2y_1 + y_2 \geq -3, \\
x_2 + 3y_1 - 4y_2 \geq 4, \\
y_1 \geq 0, y_2 \geq 0.
\end{aligned}
\tag{1}
$$

We show that the inner problem is convex, hence the KKT conditions of the inner problem are necessary and sufficient for optimality and can simple be added into the constraints of the outer problem:

$$
\begin{aligned}
\nabla_y \left[2x_1^2 + y_1^2 - 5y_2\right] &= [2y_1,\ -5]^T \\
\nabla_y^2 \left[2x_1^2 + y_1^2 - 5y_2\right] &= \begin{bmatrix} 2 & 0 \\ 0 & 0 \end{bmatrix}
\end{aligned}
$$

Since the Hessian of the inner problem is positive semi-definite we know that the inner problem in convex.

1.1 First-Order Optimality Conditions

We introduce surplus variables $(b_1,\ b_2) \geq 0$ in the inequalities of the inner problem to obtain equalities. Then, we can set up the KKT necessary conditions of the inner problem and add them as constraints to the outer problem.
As a Preliminary to stating the necessary conditions, we define the Lagrangian function as belows.

$$\mathcal{L}(x, \lambda) = f(x) - \sum_{i \in \mathcal{E} \cup \mathcal{I}} \lambda_i c_i(x). \tag{2}$$

Where, $\mathcal{E}$ and $\mathcal{I}$ are the set of equalities and inequalities, respectively. In our problem, we can define the Lagrangian function as following:

$$\begin{aligned} \mathcal{L}(x, \lambda) = 2x_1^2 + y_1^2 - 5y_2 \\ -\lambda_1(x_1^2 - 2x_1 + x_2^2 - 2y_1 + y_2 + 3 - b_1) \\ -\lambda_2(x_2 + 3y_1 - 4y_2 - 4 - b_2), \\ \text{where} \quad b_1, b_2 \geq 0 \end{aligned} \tag{3}$$

The necessary conditions defined in the following theorem are called **first-order-conditions** because they are concerned with properties of the gradients (first-derivative vectors) of the objective and constraint functions.

1.2 First-Order Necessary Conditions

Suppose that x^* is a local solution that the function f and c_i are continuously differentiable and that the LICQ holds at x^*. Then there is a Lagrange multiplier vector λ^*, with components λ_i^*, $i \in \mathcal{E} \cup \mathcal{I}$, such that the following conditions are satisfied at $(x^*,\ \lambda^*)$.

$$\nabla_x \mathcal{L}(x^*, \lambda^*) = 0, \tag{4}$$

$$c_i(x^*) = 0, \quad \forall i \in \mathcal{E}, \tag{5}$$

$$c_i(x^*) \geq 0, \quad \forall i \in \mathcal{I}, \tag{6}$$

$$\lambda_i^* \geq 0, \quad \forall i \in \mathcal{I}, \tag{7}$$

$$\lambda_i^* c_i(x^*) = 0, \quad \forall i \in \mathcal{E} \cup \mathcal{I} \tag{8}$$

These are known as the Karush-Kuhn Tucker conditions, or KKT conditions for short. Conditions (8) are know as the complementarity conditions (or, complementary slackness), they imply that either constraint i is active or $\lambda_i^* = 0$, or possibly both. In particular, when there are no inequality constraints we can eliminate conditions (4) (5) and (6). Thus the KKT conditions become just the condition (1), with the vector λ^* of **Lagrange multipliers** .

According to above necessary conditions, we can eliminate the inner problem by adding these constraints to the outer problem:

Equation (1) in the KKT conditions:

$$\nabla_{\mathbf{y}}\mathcal{L}(\mathbf{y},\lambda) = \begin{bmatrix} 2y_1 + 2\lambda_1 - 3\lambda_2 \\ -5 - \lambda_1 + 4\lambda_2 \end{bmatrix} = \mathbf{0}. \tag{9}$$

The equality constraints:

$$\begin{aligned} x_1^2 - 2x_1 + x_2^2 - 2y_1 + y_2 + 3 - b_1 &= 0 \\ x_2 + 3y_1 - 4y_2 - 4 - b_2 &= 0 \end{aligned} \tag{10}$$

And the KKT conditions (12.34 d,e):

$$\begin{aligned} \lambda_1 &\geq 0 \\ \lambda_2 &\geq 0 \\ \lambda_1(x_1^2 - 2x_1 + x_2^2 - 2y_1 + y_2 + 3) &= 0 \\ \lambda_2(x_2 + 3y_1 - 4y_2 - 4) &= 0 \end{aligned} \tag{11}$$

1.3 Sufficient Conditions

In some cases, the necessary conditions are also sufficient for optimality. In general, the necessary conditions are not sufficient for optimality and additional information is necessary, such as **Second Order Sufficient Conditions (SOSC)**. For smooth function, SOSC involve the second derivatives.
The necessary conditions are sufficient for optimality if the objective function f of maximization problem is a concave function, the inequality constraints are continuously differentiable convex functions and the equality constraints are affine functions.

1.4 Results

We submitted our AMPL model **nlp.mod** to the nonlinear constrained optimization solver Knitro with various starting points x_0, y_0. Each of our submissions returned the same optimal solution and objective function value. Finally we submitted our model to the global optimization solver Couenne to confirm our solutions:

x_0	y_0	objective	x^*	y^*
(0,0)	(0,0)	-12.679	(8.8e-4,2)	(1.875,0.906)
(1.45,.45)	(1.88,.64)	-12.679	(1.2e-4,2)	(1.875,0.906)
(5,-4)	(3,-1)	-12.679	(0,2)	(1.875,0.906)
(1.5,.7)	(1.88,.5)	-12.679	(0,2)	(1.875,0.906)
Couenne		-12.679	(0,2)	(1.875,0.906)

We are puzzled by these results because the outer objective function is not convex since the Hessian is not positive semidefinite:

$$\nabla_x F(x) = \begin{bmatrix} -2x_1 \\ -3 \end{bmatrix} \rightarrow \nabla_x^2 F(x) = \begin{bmatrix} -2 & 0 \\ 0 & 0 \end{bmatrix}$$

Thus we would expect that there would be more than one local minima. We did find any other local minima in our trials however.

2 Robust Optimization

The problem defined in terms of profit maximization of the farmer is as below:

" *A farmer has some cows and sheep. Stables are available for 50 cows and 300 sheep. The pasture is 72 acres. A cow needs 1 acre and a sheep 0.2. Labor is available for 10,000 hours per year with a cow requiring 150 hours and a sheep 25. The farmer's profit is* 250 *per cow and* 45 *per sheep.*"

Translating the problem as an LP, let us define the number of cows be x_1 and the number of sheep be x_2.

$$\max_{x_1,x_2} \quad 250x_1 + 45x_2 \tag{12}$$

$$x_1 + 0.2x_2 \leq 72 \tag{13}$$

$$150x_1 + 25x_2 \leq 10000 \tag{14}$$

$$0 \leq X \leq 50 \tag{15}$$

$$0 \leq Y \leq 300 \tag{16}$$

$$\tag{17}$$

(note that we do not restrict our solutions to be integers as would be suggested by the nature of the problem)

Solving LP: When this LP is solved using CPLEX we obtain the **maximimum profit of 17200 with 40 cows (x_1) and 160 sheep (x_2)**. The solution can be found by executing **lp.mod**.

Once we allow the data perturbation to occur we can obtain the range of predictable swings in profit of the farm. Here the perturbation parameter $\hat{a}_{ij} = 0.5A$, where $A = \begin{bmatrix} 1 & 0.2 \\ 150 & 25 \end{bmatrix}$

2.1 Soyster's Method

Given the uncertainity model of the data $\mathcal{U}$ we have the robust formulation as:

$$\text{maximize} \quad C^{'}X \tag{18}$$

$$\text{subject to} \quad \sum_j a_{ij}x_j + \sum_{j\in J_i} \hat{a}_{ij}y_j \leq b_i \quad \forall i \tag{19}$$

$$-y_j \leq x_j \leq y_j \quad \forall j \tag{20}$$

$$l \leq X \leq u \tag{21}$$

$$Y \geq 0 \tag{22}$$

$$\tag{23}$$

We ran the simulations using this LP model and generated the set of results below:

Table 1: Using Soyster's Method

Trial	Objective	$\hat{a}_{ij}$	Cows (x_1)	Sheeps (x_2)
1	17114.4	0.005A	39.80	159.20
2	16381	0.05A	38.10	152.38
3	13760	0.25A	32	128

Our main observation from this model is this type of model is the most conservative in practice leading to lower objective value than the exact objective found by solving the LP problem. One way to look at this is that the actual value of the perturbation in the data depends directly on it's noise. However, limiting it by one linear scale factor would not sufficit the entire distribution of solutions possible. Making it necessary to add additional variables into the constraints, these variables are taken from a distribution such as a Gaussian. The solution can be found by executing **lp_rob1.mod**.

2.2 Ben-Tal Nemirovski

$$\text{maximize} \quad C^{'}X \tag{24}$$

$$\text{subject to} \quad \sum_j a_{ij}x_j + \sum_{j\in J_i} \hat{a}_{ij}y_j + \Omega_i\sqrt{\sum_{j\in J_i} \hat{a}_{ij}^2 z_{ij}^2} \leq b_i \quad \forall i \tag{25}$$

$$-y_{ij} \leq x_j - z_{ij} \leq y_{ij} \quad \forall i, j \in J_i \tag{26}$$

$$l \leq X \leq u \tag{27}$$

$$Y \geq 0 \tag{28}$$

$$\tag{29}$$

Notice that the problem now has a nonlinear constraint. We submitted this AMPL code to the Knitro solver. We believe that Ben-Tal formulation is on the basis of introducing probability distributions into the LP problem to allow for

Table 2: Ben-Tal Nemirovski's Method

Trial	Objective	$\hat{a}_{ij}$	Ω_i	Cows(x_1)	Sheeps(x_2)
1	17193.8	0.05A	0.01	39.98	159.95
2	17138.6	0.05A	0.1	39.84	159.53
3	17047.2	0.05A	0.25	39.60	158.83
4	16897.2	0.05A	0.50	39.21	157.66

even slightly relaxed solutions when $\Omega_j \neq 0$. As shown in our table of results, as Ω_j increases, the optimal objective function value decreases. The solution can be found by executing **lp_rob2.mod**. Notice that in all cases Ω_i we have an objective value that is higher than in the case of Soyster's method with that same value of $\hat{a} = .05$.

Additional Experiment

We can analysis how much the objective is perturbed for the two robust models above mentioned. In order to obtain the value of perturbation (P_value), we can defined simple equation as following :

$$P\_value = \frac{|original\_LP\_objective - robust\_LP\_objective|}{|original\_LP\_objective|}. \tag{30}$$

The table 3 presents results obtaining from above equation for two robust models with different perturbations (0.5%, 5% and 25%). Note that we use omega=.1 for Ben-Tal Nemirovski

Table 3: perturbation for the two robust models

Model	P_value(0.5%)	P_value(5%)	P_value(25%)
Soyster	0.005	0.0476	0.2000
Ben-Tal	3.4884e-04	0.0035	0.0198

As you can see, the Ben-Tal method had lower perturbations from the original LP optimal objective value.

Table 4: Varying Ω of Ben-Tal and its corresponding perturbation values

Model	Ω	P_value * 100
Ben-Tal	0.01	0.036047 %
	0.1	0.35698 %
	0.25	0.88837 %
	0.5	1.7605 %

This shows that as Ω increases, so to does the perturbation from the original LP objective value.