SPEAKER: ADAM VAN TUYL

INTRODUCTION

Last class we were introduced to the notion of nonlinear programming, i.e., a mathematical programming problem where either the constraints or objective function (or both) are given by nonlinear functions. In today's class we will focus on a special subclass, namely, the *quadratic programming (QP) problems*. Quadratic programing problems are very similar to linear programming problems. The main difference between the two is that our objective function is now a quadratic function (the constraints are still linear). The goal of this talk is to introduce quadratic programming problems, and a variation of the simplex method that can be used to solve these problems when the objective function is concave. The material of this talk is based upon [1, Section 14.7]. The example that I used is [1, Exercise 44, pg. 572]. I also found the notes of Jensen and Bard [2] to be helpful.

QUADRATIC PROGRAMMING

Quadratic Programs (QP) have the form

$$\begin{array}{lrcl} \text{Maximize:} & \mathbf{c}\mathbf{x} - \frac{1}{2}\mathbf{x}^T Q\mathbf{x} & = & f(\mathbf{x}) \\ \text{Subject to:} & A\mathbf{x} & \leq & \mathbf{b} \\ & \mathbf{x} & \geq & \mathbf{0} \end{array}$$

where $\mathbf{c}$ is a row vector, $\mathbf{x}$ and $\mathbf{b}$ are column vectors, and A and Q are matrices.

A few words on the matrix Q. The matrix Q is chosen so that Q is an $n \times n$ symmetric matrix , i.e., $q_{i,j} = q_{j,i}$ for all i, j. Finding the matrix Q is described in many linear algebra books under the topic of quadratic forms. We illustrate with an example.

Example 1. Consider the quadratic function

$$f(x_1, x_2, x_3) = 15x_1 + 30x_2 + 17x_3 + 5x_1^2 + 3x_2^2 + 2x_3^2 - x_1x_2 + 8x_2x_3.$$

The linear part is simply given by

$$\begin{bmatrix} 15 & 30 & 17 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = 15x_1 + 30x_2 + 17x_3.$$

To make the matrix Q, we make a 3×3 matrix where in positions $q_{i,j} = q_{j,i}$ we put the negative of the coefficient of x_ix_j. (If $i = j$, $q_{i,i}$ is -2 times the coefficient of x_i^2.) So, in our example

$$Q = \begin{bmatrix} -10 & 1 & 0 \\ 1 & -6 & -8 \\ 0 & -8 & -4 \end{bmatrix}.$$

So,

$$\begin{aligned}
-\frac{1}{2}\mathbf{x}^T Q\mathbf{x} &= -\frac{1}{2}\begin{bmatrix} x_1 & x_2 & x_3 \end{bmatrix}\begin{bmatrix} -10 & 1 & 0 \\ 1 & -6 & -8 \\ 0 & -8 & -4 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \\
&= -\frac{1}{2}(-10x_1^2 + 2x_1x_2 - 6x_2^2 - 16x_2x_3 - 4x_3^2) \\
&= 5x_1^2 + 3x_2^2 + 2x_3^2 - x_1x_2 + 8x_2x_3.
\end{aligned}$$

The method that we will describe will give a global maximum provided that our objective function is *concave* for all $\mathbf{x}$.

Definition 2. A symmetric matrix Q is *semi-positive definite* if $\mathbf{x}^T Q\mathbf{x} > \mathbf{0}$ for all $\mathbf{x} \geq 0$.

We have the following theorem:

Theorem 3. *The objective function*

$$f(\mathbf{x}) = \mathbf{c}\mathbf{x} - \frac{1}{2}\mathbf{x}^T Q\mathbf{x}$$

is concave if Q is semi-positive definite. Moreover, Q is semi-positive definite if and only if all the eigenvalues of Q are nonnegative.

KTT Conditions applied to QP

We begin with a result independently due to Karush and to Kuhn and Tucker, usually called the KTT conditions. The KTT conditions describe some of the properties that an optimal solution to a nonlinear program must satisfy.

Theorem 4. *Suppose that we have a nonlinear program of the form*

$$\textit{maximize } f(\mathbf{x})$$

subject to the constraints $g_i(\mathbf{x}) \leq b_i$ for $i = 1, \ldots, m$ and $\mathbf{x} \geq \mathbf{0}$. Furthermore, assume that f and the g_i's are differentiable. If $\mathbf{x}^\star = (x_1^\star, \ldots, x_n^\star)$ is an optimal solution, then there exists m numbers $u_1, \ldots, u_m$ such that all the following conditions are satisfied:

1. *At $\mathbf{x}^\star$*
$$\frac{\partial f}{\partial x_j} - \sum_{i=1}^{m} u_i \frac{\partial g_i}{\partial x_j} \leq 0 \text{ for } j = 1, \ldots, n.$$
2. $$x_j^\star \left(\frac{\partial f}{\partial x_j} - \sum_{i=1}^{m} u_i \frac{\partial g_i}{\partial x_j} \right) = 0 \text{ for } j = 1, \ldots, n.$$
3. $g_i(\mathbf{x}^\star) - b_i \leq 0$ *for* $i = 1, \ldots, m$.
4. $u_i(g_i(\mathbf{x}^\star) - b_i) = 0$ *for* $i = 1, \ldots, m$.
5. $x_j^\star \geq 0$ *for* $j = 1, \ldots, n$.
6. $u_i \geq 0$ *for* $i = 1, \ldots, m$.

Note that this describes the properties of the optimal solution, but it in general, this list does not completely describe the optimal solution. However, in some cases, the above list completely describes the optimal solution.

Corollary 5. *Let $f(\mathbf{x})$ be a concave function, and suppose that the constraints $g_i(\mathbf{x})$ are convex functions for all i. Then $\mathbf{x}^\star = (x_1^\star, \ldots, x_n^\star)$ is an optimal solution if and only if all of the conditions $1-6$ of Theorem 4 are satisfied.*

Before working out the details of a specific example, we sketch out our approach to solving a QP. We begin with a QP whose objective function is concave. We can use Theorem 3 to verify this condition. Since linear constraints are convex, we can apply Corollary 5. In particular, we want to solve conditions 1-6 of Theorem 4 when we start with a QP. Note that when we take derivatives, we are going to get a series of linear equations (almost! there are some subtleties), and we can then solve these equations using a modified version of the simplex method.

We will now discuss a specific example for the remainder of the talk.

Example 6. We want to solve the following QP:

$$\begin{array}{lrcl} \text{Maximize:} & 8x_1 - x_1^2 + 4x_2 - x_2^2 & = & f(x_1, x_2) \\ \text{Subject to:} & x_1 + x_2 & \leq & 2 \\ & x_1, x_2 & \geq & 0. \end{array}$$

Rewriting this equation, we get

$$f(x_1, x_2) = \begin{bmatrix} 8 & 4 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} - \frac{1}{2} \begin{bmatrix} x_1 & x_2 \end{bmatrix} \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}.$$

The matrix $Q = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}$ is semi-positive definite since its only eigenvalue is 2.

By apply the KTT conditions, we are looking for numbers x_1, x_2 and u_1 such that

$$\begin{array}{rcl} 8 - 2x_1 - u_1 & \leq & 0 \\ 4 - 2x_2 - u_1 & \leq & 0 \\ x_1(8 - 2x_1 - u_1) & = & 0 \\ x_2(4 - 2x_2 - u_1) & = & 0 \\ x_1 + x_2 - 2 & \leq & 0 \\ u_1(x_1 + x_2 - 2) & = & 0 \\ x_1, x_2, u_1 & \geq & 0. \end{array}$$

We add slack variables y_1, y_2 and v_1 for each of three inequalities as follows, and after rearranging, we get:

$$\begin{aligned} -2x_1 - u_1 + y_1 &= -8 \\ -2x_2 - u_1 + y_2 &= -4 \\ x_1(8 - 2x_1 - u_1) &= 0 \\ x_2(4 - 2x_2 - u_1) &= 0 \\ x_1 + x_2 + v_1 &= 2 \\ u_1(x_1 + x_2 - 2) &= 0 \\ x_1, x_2, u_1 &\geq 0. \end{aligned}$$

Since $y_1 = -8 + 2x_1 + u_1$, the third equation can be rewritten as $x_1 y_1 = 0$ (technically, we have $x_1(-y_1) = 0$, but note that the sign doesn't change the fact that the third equation is true if and only if either $x_1 = 0$ or $y_1 = 0$). Similarly, the fourth equation becomes $x_2 y_2$ and the sixth equation becomes $u_1 v_1 = 0$. The pairs $(x_1, y_1), (x_2, y_2)$ and (u_1, v_1) are the *complementary variables.* We can combine the third, fourth, and sixth equations into one constraint

$$x_1 y_1 + x_2 y_2 + u_1 v_1 = 0$$

which is called the *complementary constraint.* Given a pair of complementary variables, at most one can be nonzero.

After rearrangement, our set of conditions become:

$$\begin{aligned} 2x_1 + u_1 - y_1 &= 8 \\ 2x_2 + u_1 - y_2 &= 4 \\ x_1 + x_2 + v_1 &= 2 \\ x_1 y_1 + x_2 y_2 + u_1 v_1 &= 0 \\ x_1, x_2, u_1, y_1, y_2, v_1 &\geq 0. \end{aligned}$$

These look almost like linear constraints, except for the fourth condition.

We pause from our example to summarize the above procedure for the general case.

Theorem 7. *Given any quadratic program of the form*

$$\begin{aligned} \textit{Maximize:} \quad \mathbf{c}\mathbf{x} - \tfrac{1}{2}\mathbf{x}^T Q\mathbf{x} &= f(\mathbf{x}) \\ \textit{Subject to:} \quad A\mathbf{x} &\leq \mathbf{b} \\ \mathbf{x} &\geq \mathbf{0}, \end{aligned}$$

the KTT conditions can be expressed as

$$\begin{aligned} Q\mathbf{x} + A^T\mathbf{u} - \mathbf{y} &= \mathbf{c}^T \\ A\mathbf{x} + \mathbf{v} &= \mathbf{b} \\ \mathbf{x}, \mathbf{u}, \mathbf{v}, \mathbf{y} &\geq \mathbf{0} \ \textit{(of the appropriate size)} \\ \mathbf{x}^T\mathbf{y} + \mathbf{u}^T\mathbf{v} &= \mathbf{0} \end{aligned}$$

where $\mathbf{u}$ *is the column vector of* u_i*'s and* $\mathbf{v}$ *and* $\mathbf{y}$ *are the column vectors of slack variables.*

Example 8. (Continued) Our next step is to find any solution to the system of equations given before Theorem 7. By Corollary 5, such a solution will correspond to an optimal solution of our original QP. We therefore want to get our hands on one solution to this system.

Our strategy is to use the simplex method to find such a solution. We will view the equations as the linear constraints of some linear programming problem. The role of the complementary constraint is to change our criterion for deciding the entry variable in the simplex method.

To turn this problem into a linear programming problem, we need an objective function. We introduce *artificial variables* to each equation such that $c_i > 0$ or $b_j < 0$.[1] In our problem, we get

$$\begin{aligned} 2x_1 + u_1 - y_1 + z_1 &= 8 \\ 2x_2 + u_1 - y_2 + z_2 &= 4 \\ x_1 + x_2 + v_1 &= 2 \\ x_1y_1 + x_2y_2 + u_1v_1 &= 0 \\ x_1, x_2, u_1, y_1, y_2, v_1, z_1, z_2 &\geq 0. \end{aligned}$$

Note that if set all the variables except z_1 and z_2 to zero, then we get a solution to the above system, namely $z_1 = 8$, $z_2 = 4$, and all other variables are equal to zero. We, however, want a solution where $z_1 = z_2 = 0$. So, we want to move to an alternative basic solution where z_1 and z_2 become non-basic variables. Notice that such a solution will also satisfy our original system (i.e., the equations before Theorem 7).

We then want to minimize the equation

$$\text{Minimize } Z = z_1 + z_2.$$

To turn this into a linear programming problem we can solve with the simplex method, we turn it into a maximizing problem, i.e., we want to

$$\text{Maximize } Z' = (-Z) = -z_1 - z_2.$$

By rearranging our linear conditions, we get

$$Z = (8 - 2x_1 - u_1 + y_1) + (4 - 2x_2 - u_1 + y_2) = 12 - 2x_1 - 2x_2 - 2u_1 + y_1 + y_2.$$

So, to solve our original quadratic program, we use the simplex method to solve the following linear program:

$$\begin{aligned} \text{Maximize:} \quad -12 + 2x_1 + 2x_2 + 2u_1 - y_1 - y_2 &= Z' \\ \text{Subject to:} \quad x_1 + u_1 - y_1 + z_1 &= 8 \\ 2x_2 + u_1 - y_2 + z_2 &= 4 \\ x_1 + x_2 + v_1 &= 2 \\ x_1, x_2, u_1, y_1, y_2, v_1, z_1, z_2 &\geq 0. \end{aligned}$$

But before you go off and do that, we need to explain how to handle the constraint:

$$x_1y_1 + x_2y_2 + u_1v_1 = 0.$$

[1] In some textbooks, an artificial variable is introduced for every linear equation. Although both methods are probably equivalent, I couldn't find a proof of this fact. I will follow the method described in [1].

Note that if in our simplex method, if x_1 is already chosen as a basic variable, then we cannot chose y_1 to also be a basic variable. Indeed, if they were both basic, then (in the majority of cases), x_1 and y_1 would be nonzero, contradicting the above constraint equation. So, we need to change the entry rule (i.e., deciding which column is a pivot column). Precisely, we have

> **New Entry Rule.** When picking a new basic variable (i.e., pivot column), eliminate for consideration any variable that is the complementary variable of any variable that is currently a basic variable.

For completeness, we finish our example. Our initial tableau is

x_1	x_2	u_1	y_1	y_2	v_1	z_1	z_2	
2	0	1	−1	0	0	1	0	8
0	2	1	0	−1	0	0	1	4
1	1	0	0	0	1	0	0	2
−2	−2	−2	1	1	0	0	0	−12

Our initial basic variables are (v_1, z_1, z_2) with initial feasible solution $(0, 0, 0, 0, 0, 0, 2, 8, 4)$.

We pick x_1 as our new entering variable (this is okay with new entry rule since y_1 is not a basic variable). The third row is our new pivot column, so v_1 is a leaving variable. After this iteration, our tableau looks like:

x_1	x_2	u_1	y_1	y_2	v_1	z_1	z_2	
0	−2	1	−1	0	−2	1	0	4
0	2	1	0	−1	0	0	1	4
1	1	0	0	0	1	0	0	2
0	0	−2	1	1	2	0	0	−8

Our basic variables are (x_1, z_1, z_2) with basic solution $(2, 0, 0, 0, 0, 0, 4, 4)$.

At the next iteration, y_1 is eliminated as a candidate for a basic variable (in this situation, we wouldn't even consider y_1 since the number in the corresponding column of the objective row is positive). Our only choice for a basic variable is u_1. Our leaving variable is z_1 (you could also choose z_2 since there is a tie). After pivoting, we get:

x_1	x_2	u_1	y_1	y_2	v_1	z_1	z_2	
0	−2	1	−1	0	−2	1	0	4
0	4	0	1	−1	2	−1	1	0
1	1	0	0	0	1	0	0	2
0	−4	0	−1	1	−2	2	0	0

Our basic variables are (x_1, u_1, z_2) with basic solution $(2, 0, 4, 0, 0, 0, 0, 0)$.

At our next iteration, we get x_2 is our new basic variable, and z_2 is our leaving variable. Pivoting gives:

x_1	x_2	u_1	y_1	y_2	v_1	z_1	z_2	
0	0	1	$-\frac{1}{2}$	$-\frac{1}{2}$	-1	$\frac{1}{2}$	$\frac{1}{2}$	4
0	1	0	$\frac{1}{4}$	$-\frac{1}{4}$	$\frac{1}{2}$	$-\frac{1}{4}$	$\frac{1}{4}$	0
1	0	0	$-\frac{1}{4}$	$\frac{1}{4}$	$\frac{1}{2}$	$\frac{1}{4}$	$-\frac{1}{4}$	2
0	0	0	0	0	0	1	1	0

So, our basic variables are (x_1, x_2, u_1) with feasible solution $(2, 0, 4, 0, 0, 0, 0, 0)$. This is our optimal solution since our simplex method has completed.

Notice that until the last step, we had either z_1 or z_2 as a basic variable. However, at the last step, neither variable is a basic variable, i.e., $z_1 = z_2 = 0$. So, our feasible solution is also a solution to our system of equations given before Theorem 7. By Corollary 5, this solution now gives an optimal solution to the original QP. In particular, $(x_1, x_2) = (2, 0)$ gives the optimal solution.

If, on the other hand, the variable z_1 or z_2 must always be a basic variable, then there is no feasible solution only in $x_1, x_2, u_1, y_1, y_2, v_1$. In this case there is no solution to the QP.

Final Comments

At the heart of QP is the KTT conditions. Note that these conditions hold true for many nonlinear integer programs. QP have many applications in economics, and is sometimes considered a separate sub-discipline. A quick search of the web produces many alternative means to solve a QP. In many cases, software can be downloaded and played with.

Problems from Lecture 7

1. Use the method described in this lecture to solve the QP:

$$\begin{array}{lrcl} \text{Maximize:} & 15x_1 + 30x_2 + 4x_1x_2 - 2x_1^2 - 4x_2^2 & & \\ \text{Subject to:} & x_1 + 2x_2 & \leq & 30 \\ & x_1, x_2 & \geq & 0 \end{array}$$

   Hint: This is the example discussed in [1, Section 14.7].

References

[1] F.S. Hillier, G.J. Lieberman, *Introduction to Mathematical Programming.* McGraw-Hill, Toronto, 1990.
[2] P.A. Jensen, J.F. Bard, Lecture Notes on Quadratic Programming. `https://www.me.utexas.edu/~jensen/ORMN/supplements/methods/nlpmethdo/nlp_intro.html`