Quadratic Optimization

This handout covers some results on constrained nonlinear optimization. Suppose x is a variables vector of n dimensions: $x = (x_1, \ldots, x_n)$. We consider simple polynomial function of x that involve at most second-order terms. Matrix and vector notation considerably simplify the expressions. Here is a reminder of several formulas, in terms of the vectors x, y and the matrices A, B:

The inner product of x and y	$x'y = x_1y_1 + \ldots + x_ny_n$
The squared norm of x	$\lvert x\rvert^2 = x'x = x_1^2 + \ldots + x_n^2$
The most general quadratic form	$x'Ax + b'x + c$
Remainder	$x'Ax = x'(Ax) = \sum_{ij} a_{ij}x_ix_j$
Positive semidefinite matrix B	$x'Bx \geq 0 \quad$ whenever $x \neq 0$
Gradients (derivatives)	$f(x) = x'Bx + b'x + c,\ B$ pos semidef, then $\nabla f(x) = 2Bx + b$

Unconstrained optimization

Here we would like to compute the minimum of $f(x)$. It can be shown that a quadratic $f(x) = x'Bx+b'x+c$ has a minimum if and only if B is positive semidefinite. It has a unique minimum if and only if B is positive definite. In both cases the minimum is values of x that zero out the gradient. That is, the solutions of:

$$Bx = -\frac{b}{2}$$

Example 1

Find the minimum of

$$f(x_1, x_2) = x_1^2 + 2x_2^2 + 2x_1x_2 + 2x_1 - 6x_2 + 4$$

With $x = (x_1, x_2)$ we have:

$$f(x) = x' \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix} x + (2, -6)x + 4$$

Therefore, if there is a minimum it is the solution of:

$$\begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix} x = -(2, -6)/2 = (-1, 3)$$

The solution is $(-5, 4)$ so that if a minimum exists it is at $x_1 = -5, x_2 = 4$. To argue that it is indeed a minimum we still need to show that the matrix is positive definite. One way of showing that a matrix B is positive semidefinite is to show that there is a matrix G such that:

$$B = G'G$$

In our case with $B = \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix}$ we can take $G = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$.

Equality linear constraints

Here we would like to compute the minimum of $f(x)$ where x is also subject to linear equality constraints. A linear equality constraint can be written as $c'x = d$ where c is a vector and d is a scalar. When there are k constraints we write them as $c_i'x = d_i$.

There are several methods to reduce this to the case of unconstrained optimization. The most basic one is the method of Lagrange multipliers. In order to minimize $f(x)$ under the above equality constraints we define a Lagrangian multiplier α_i for the ith constraint and form the Lagrangian function $L(x, \alpha_1, \ldots, \alpha_k)$ as follows:

$$L(x, \alpha_1, \ldots, \alpha_k) = f(x) + \sum_{i=1}^{k} \alpha_i (c_i'x - d_i)$$

We can now treat L as a function of $n + k$ variables and use the method of unconstrained minimization to find the unconstrained minimum of L. (Observe that L is quadratic.)

Example 2

Minimize $f(x_1, x_2, x_3) = x_1^2 + x_2^2 + x_3^2$ under the two constraints: $x_2 - x_1 = 1$, $x_3 = 1$.
This is an easy case to solve without the Lagrange technique. It is easy to see by direct substitution that $f = x_1^2 + (x_1 + 1)^2 + 1$ and by taking derivatives and equating to 0 the minimum is at $x_1 = -1/2$. Using the Lagrange technique:

$$L(x_1, x_2, x_3, \alpha_1, \alpha_2) = x_1^2 + x_2^2 + x_3^2 + \alpha_1(x_2 - x_1 - 1) + \alpha_2(x_3 - 1)$$

The derivative with respect to x_1 gives:	$2x_1 - \alpha_1 = 0$
The derivative with respect to x_2 gives:	$2x_2 + \alpha_1 = 0$
The derivative with respect to x_3 gives:	$2x_3 + \alpha_2 = 0$
The derivative with respect to α_1 gives:	$x_2 - x_1 - 1 = 0$
The derivative with respect to α_2 gives:	$x_3 - 1 = 0$

Treating this as a system of 5 equations with 5 unknowns we get the following solution:

$$x_1 = -1/2, \quad x_2 = 1/2, \quad , x_3 = 1, \quad \alpha_1 = -1, \quad \alpha_2 = -2$$

The dual problem

There is another way of using the Lagrangian to solve the constrained optimization problem. The following result can be proved:

$$\min_{x \text{ subject to the equality constraints}} f(x) = \max_{\alpha_1, \ldots, \alpha_k} \min_x L(x, \alpha_1, \ldots, \alpha_k)$$

To use this, we begin by treating the Lagrangian multipliers as constants, minimizing the Lagrangian with respect to x. This is done by taking the derivatives only with respect to x. The resulting system of linear equations can be used to solve for the variables x in terms of the Lagrangian multipliers. This is then substituted back into the Lagrangian, transforming it into a function of only the Lagrangian multipliers. This function is called **the dual problem**, and it can then be solved using unconstrained techniques.

Example 3

To convert the problem of Example 2 into a dual form we only need to take the derivatives of the Lagrangian with respect to x_1, x_2, x_3 As above:

The derivative with respect to x_1 gives:	$2x_1 - \alpha_1 = 0$
The derivative with respect to x_2 gives:	$2x_2 + \alpha_1 = 0$
The derivative with respect to x_3 gives:	$2x_3 + \alpha_2 = 0$

Solving for x_1, x_2, x_3 in terms of α_1, α_2 we have:

$$x_1 = \alpha_1/2, \quad x_2 = -\alpha_1/2, \quad x_3 = -\alpha_2/2 \tag{1}$$

Substituting back into the Lagrangian gives:

$$L(\alpha_1, \alpha_2) = -\frac{1}{2}\alpha_1^2 - \frac{1}{4}\alpha_2^2 - \alpha_1 - \alpha_2$$

The dual problem requires that we maximize this expression The maximizing α_1, α_2 are: $\alpha_1 = -1, \alpha_2 = -2$, and the minimizing x_1, x_2, x_3 of the original (primal) problem are computed from (1).

Example 4

Consider the n dimensional hyperplane (generalization of a line in 2D) $w'x = s$. In 2D it is the line $w_1x_1 + w_2x_2 = s$. Compute the distance of this hyperplane from the origin.
A point x on the hyperplane has distance of $|x|$ from the origin. We are looking for the point x that minimizes $|x|^2$ subject to the constraint $w'x = s$. The Lagrangian:

$$L(x, \alpha) = |x|^2 + \alpha(w'x - s)$$

Computing the derivative with respect to x and equating to 0 we have:

$$2x + \alpha w = 0 \qquad \Rightarrow \quad x = -\frac{\alpha}{2}w$$

Substituting into the Lagrangian we get the dual problem:

$$L(\alpha) = \frac{\alpha^2}{4}|w|^2 + \alpha(-\frac{\alpha}{2}w'w - s) = -\frac{\alpha^2}{4}|w|^2 - \alpha s$$

In terms of α this is a parabola with maximum at $\alpha = -2s/|w|^2$. This gives: $x = sw/|w|^2$ so that the distance of the hyperplane from the origin is $|s|/|w|$.

Linear inequality constraints

Here we would like to compute the minimum of $f(x)$ where x is subject to linear inequality constraints. A linear inequality constraint can be written as $c'x \leq d$ where c is a vector and d is a scalar. When there are k constraints we write them as $c_i'x \leq d_i$.
The Lagrangian in this case is the same as in the previous case:

$$L(x, \alpha_1, \ldots, \alpha_k) = f(x) + \sum_{i=1}^{k} \alpha_i (c_i'x - d_i)$$

And the following dual theorem holds:

$$\min_{x \text{ subject to the inequality constraints}} f(x) = \max_{\alpha_1,\ldots,\alpha_k \text{ subject to } \alpha_j \geq 0} \min_x L(x, \alpha_1, \ldots, \alpha_k)$$

Formally, the problem

$$\min_{x \text{ subject to the inequality constraints}} f(x)$$

is called **the primal problem** and

$$\max_{\alpha_1,\ldots,\alpha_k \text{ subject to } \alpha_j \geq 0} \min_x L(x, \alpha_1, \ldots, \alpha_k)$$

is called **the dual problem**. Observe that here the dual problem is no longer trivial.
The following important result can be proved for the solution to the dual problem:

The Karush-Kuhn-Tucker complementarity condition:

$$\alpha_i(c_i'x - d_i) = 0 \quad \text{for} \quad i = 1, \ldots, k$$

This means that either the constraint holds as an **equality** constraint, or the corresponding α_i is 0.

Example 5

Minimize x^2 subject to $x \geq 1$ and $x \geq 2$. The minimizing value of x is clearly $x = 2$, with the first constraint inactive. We show how this can be deduced from the math describe above. First write the constraints as required:

$$-x + 1 \leq 0$$
$$-x + 2 \leq 0$$

The Lagrangian:

$$L(x, \alpha_1, \alpha_2) = x^2 + \alpha_1(-x + 1) + \alpha_2(-x + 2)$$

Now compute the dual problem. Taking the derivative w.r.t. x and equating to 0:

$$2x - \alpha_1 - \alpha_2 = 0 \qquad \Rightarrow \qquad x = \frac{\alpha_1 + \alpha_2}{2}$$

Substituting back into the Lagrangian:

$$L(\alpha_1, \alpha_2) = \frac{(\alpha_1 + \alpha_2)^2}{4} + \alpha_1(1 - \frac{\alpha_1 + \alpha_2}{2}) + \alpha_2(2 - \frac{\alpha_1 + \alpha_2}{2}) = \alpha_1 + 2\alpha_2 - \frac{(\alpha_1 + \alpha_2)^2}{4}$$

The dual problem:

$$\max_{\alpha_1 \geq 0,\ \alpha_2 \geq 0} \alpha_1 + 2\alpha_2 - \frac{(\alpha_1 + \alpha_2)^2}{4}$$

The solution is $\alpha_1 = 0,\ \alpha_2 = 4$.
To see why, assume that the optimum is obtained for $\alpha_1 = a_1 > 0, \alpha_2 = a_2 \geq 0$. Then taking $\alpha_1 = 0, \alpha_2 = a1 + a2$ increases the dual, contradiction. Now if $\alpha_1 = 0$ the global max is at $\alpha_2 = 4$.

The dual problem is assumed to be solved by a "black box" software.