Parabola Volume 34, Issue 2 (1998)

AN INTRODUCTION TO SEMI-DEFINITE OPTIMIZATION

by

Michael J Nealon[1]

Semi-definite optimization is a very new topic in the branch of applied mathematics known as *optimization*. Optimization refers to the process of finding the best way of achieving a goal. Types of practical applications for optimization include portfolio management (producing the largest profit from some investments; used in finance), design optimization (finding the cheapest and/or strongest design; used in engineering), and scheduling (minimizing time and effort wastage; used in management). The job of the mathematician is to ignore all the circumstantial information involved and to model the problem mathematically. One such model commonly arrived at is *linear programming* which is discussed in a Year 9 lobe; linear programming is a special case of semi-definite optimization. The major features of optimization will now be introduced via the simple example of linear programming. Once the ground work has been laid semi-definite optimization will be introduced.

Optimization

An optimization problem can be viewed as having three major components. The first are the *solution variables*. These represent the features of the problem that we can change, such as the amount to be invested in each scheme for a portfolio management problem. In real life applications there are typically several hundred of these variables to consider. For the purposes of this introduction only problems of two solution variables will be looked at. The other two components of an optimization problem are the constraints and the objective

[1]Michael is currently an honours student in Applied Mathematics at UNSW. The work described here was carried out as part of a vacation scholarship project supervised by Dr V. Jeyakumar.

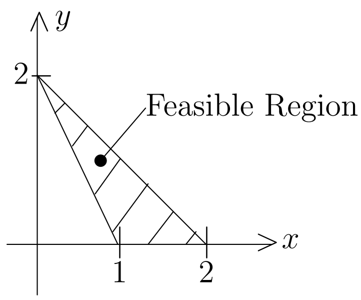
function. These components will now be looked at in detail using the linear programming example.

A *constraint* is a requirement that the optimal solution must satisfy. For example, in the scheduling problem, we require that the optimal schedule fulfils the demand. In the linear programming problem (for two variables) the constraints have the form $ax + by \leq c$. A problem of this type can have several such constraints each of which must be satisfied by the optimal solution (x^*, y^*). The region in the number plane that satisfies all these requirements is called the feasible region, as in the following example:

Example 1. Consider a linear programming problem with the following constraints:

$$
\begin{aligned}
x + y &\leq 2 \\
-2x - y &\leq -2 \\
-y &\leq 0.
\end{aligned}
$$

The feasible region is given by the following diagram:

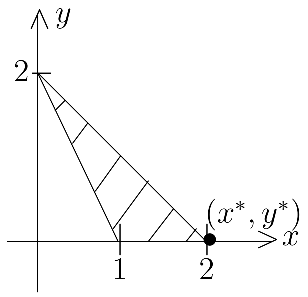


The *objective* is the function of the solution variables that we wish to optimize (usually minimize). For example in the scheduling problem this would be the function that gives time wastage in terms of the solution variables. In the case of linear programming it is a function of the form $cx + dy$. The problem is complete when a point (x^*, y^*) is found that gives the smallest possible value for $cx^* + dy^*$ under the requirement that (x^*, y^*) satisfies all the given constraints. To demonstrate, consider the following linear programming problem:

Example 2. The constraints from the previous example will be used together with an objective of $-2x + y$. This problem can be summarized with:

$$\begin{aligned} \text{minimize} \quad & -2x + y \\ \text{subject to} \quad & x + y \leq 2 \\ & -2x - y \leq -2 \\ & -y \leq 0. \end{aligned}$$

It is not difficult to see that the smallest possible value of the objective is obtained at $(2, 0)$, as shown below:



Semi-definite Optimization

In the section above the major features of optimization problems were discussed. In essence any optimization problem can be generated by varying the way in which the constraints and objective are defined. The focus will now turn to one particular problem configuration, that of semi-definite optimization. For simplicity, only problems of one solution variable (x) will be considered.

The *constraint* for the semi-definite optimization problem will be considered first. Most of the mathematics here comes directly from the 2 unit topic 'Quadratic Functions'. In particular, the definition of positive (semi)definite is required to understand the constraint. Given any quadratic $p(t) = at^2 + 2bt + c$, we say that it is positive definite if the value of $p(t)$ is positive for all values of t. The quadratic is referred to as positive semi-definite if $p(t) \geq 0$ for all values of t. Clearly $p(t) = 0$ for one value (the vertex). Algebraically, this translates to:

$p(t)$ is positive definite if:

$$a > 0$$
$$(2b)^2 - 4ac < 0$$

$p(t)$ is positive semi-definite if:

$$a > 0$$
$$(2b)^2 - 4ac = 0.$$

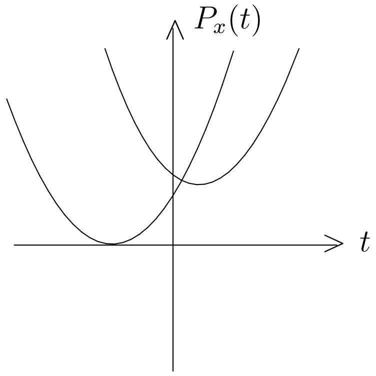
In the case of the semi-definite optimization problem the constraint is the requirement that a quadratic function is *at least* positive semi-definite. That is x (the solution variable) is allowable if

$$P_x(t) = (Ax + \alpha)t^2 + 2(Bx + \beta)t + (Cx + \gamma) \geq 0 \quad \text{for all } t.$$

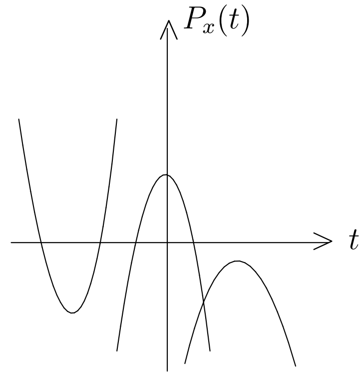
Note here we have let:

$$a = a(x) = Ax + \alpha \quad b = b(x) = Bx + \beta \quad c = c(x) = Cx + \gamma.$$

That is the parabola defined by $P_x(t)$ has to be above the t axis always (it is allowed to just touch the t axis at the vertex). The following two diagrams shows some allowable and non-allowable quadratic functions:



Allowable



Not allowable

The condition that $P_x(t) \geq 0$ for all t is equivalent to:

(1) $a > 0$

$\Rightarrow Ax + \alpha > 0$ and

(2) $4b^2 - 4ac \leq 0$

$\Rightarrow ac - b^2 \geq 0$

$\Rightarrow (Ax + \alpha)(Cx + \gamma) - (Bx + \beta)^2 \geq 0.$

The most important feature to notice about conditions (1) and (2) is that they do not depend on t, in fact the entire problem does not depend on t, which is referred to as a *dummy variable.*

We will now look at which values of x satisfy the conditions above. The allowable values depend on the constants given in $a(x)$, $b(x)$ and $c(x)$ (i.e. A, B, C, α, β, and γ). The set of all these values of x is called the *feasible set* for the problem. The optimal value of x will be one of these points.

The feasible set will now be shown to be *convex.* This property, for the one variable case, states that the set is a single interval. To have this property the feasible set must satisfy: if x and y are feasible then so is any point z between them. z will be characterized by $z = \lambda x + (1 - \lambda)y$ for any $0 \leq \lambda \leq 1$. If x and y are feasible then

$$P_x(t) = a(x)t^2 + 2b(x)t + c(x) \geq 0 \quad \text{for all } t,$$

and

$$P_y(t) = a(y)t^2 + 2b(y)t + c(y) \geq 0 \quad \text{for all } t.$$

Then

$$\begin{aligned} P_z(t) &= a(\lambda x + (1 - \lambda)y)t^2 + 2b(\lambda x + (1 - \lambda)y)t + c(\lambda x + (1 - \lambda)y) \\ &= \lambda[a(x)t^2 + 2b(x)t + c(x)] + (1 - \lambda)[a(y)t^2 + 2b(y)t + c(y)] \\ &\geq 0 \quad \text{for all } t, \text{ as } x \text{ and } y \text{ are feasible and } \lambda, (1 - \lambda) \geq 0 \end{aligned}$$

where simplifications such as:

$$\begin{aligned} a(\lambda x + (1 - \lambda)y) &= A(\lambda x + (1 - \lambda)y) + \alpha \\ &= \lambda Ax + (1 - \lambda)Ay + \alpha + \lambda\alpha - \lambda\alpha \\ &= \lambda(Ax + \alpha) + (1 - \lambda)(Ay + \alpha) \end{aligned}$$

have been used, and hence the intermediate point z is also feasible. To demonstrate the significance of this property consider the following two cases:

The case on the left is convex as any point z between feasible points x and y is feasible. The case on the right is not convex as the intermediate point z is not feasible.

Now that the constraint (and feasible set) has been defined, the focus will switch to the objective function. For the semi-definite optimization case it is simply a multiple of x (i.e. mx for some given m). The size of m does not matter, only whether it is positive or negative. We have two cases:

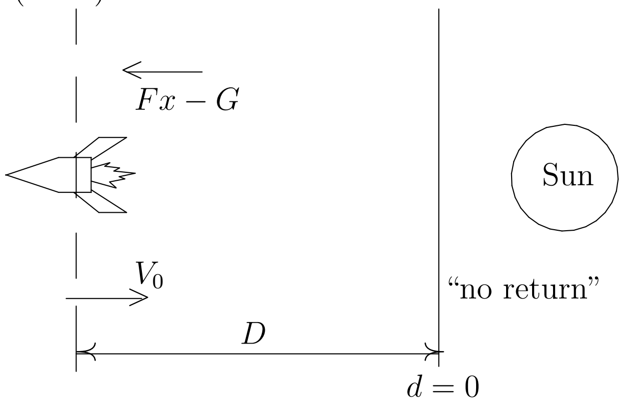
A If m is positive then minimizing mx is the same as choosing the smallest possible x. This corresponds to picking the x at the lower end of the interval feasible set.

B If m is negative then minimizing mx is the same as choosing the largest possible x. This corresponds to picking the x at the higher end of the interval feasible set.

where x^* is the optimal value of x. In the cases above two important assumptions have been made. First, the existence of a feasible x has been made. It is possible for no such x to exist (say, $A = 0$ and $\alpha < 0$.) In this case the problem is termed *infeasible*. Second, it has been assumed that the feasible set is bounded (i.e. the interval does not go to infinity in either direction). Obviously, if the first assumption fails then there is no optimal solution. If the second assumption fails, in the appropriate direction, then the optimal objective value is $-\infty$ (infinitely large in the negative direction).

Finding the solution to the problem consists of solving the inequalities (1) and (2) for the feasible set, then choosing the appropriate end-point for the optimal solution. The following example demonstrates the main features.

Example 3. Consider a rocket ship travelling through space on a long journey back to Earth. Along the way the pilot (a mathematician at heart) finds herself trapped in the Sun's gravitational pull. Before she puts the ship's engine into action she gets out her note book and makes some calculations. She is very low on fuel and must conserve as much as possible for the journey home, she therefore decides to *minimize* the rate fuel is used in this escape exercise. She also calculates the closest point his ship can get to the Sun without frying it (this point will be taken as $d = 0$). For simplicity the acceleration due to the Sun's pull will be taken to be the constant value G. The pilot draws the following diagram to represent the initial situation ($t = 0$):



Here V_0 is the initial velocity of the ship towards the Sun, D is the initial displacement from the point of "no return", and F relates fuel consumption to acceleration. The variable x is the constant rate of fuel to be used throughout the escape, so $\ddot{d} = Fx - G$ is the constant acceleration over all time. Next the pilot uses the standard process to find displacement as function of time t:

$$\ddot{d} = Fx - G$$
$$\dot{d} = (Fx - G)t - V_0$$
$$d = \frac{1}{2}(Fx - G)t^2 - V_0 t + D.$$

Clearly, the *constraint* is to keep $d \geq 0$ for all $t \geq 0$, however since V_0, $D \geq 0$ then this is the same as $d \geq 0$ for all t. The pilot realises this is a standard semi-definite optimization problem, that is, minimize x subject to the constraint that $\frac{1}{2}(Fx - G)t^2 - V_0 t + D \geq 0$ for all t. To clarify this the constants are:

$$A = \frac{1}{2}F \qquad \alpha = -\frac{1}{2}G$$
$$B = 0 \qquad \beta = -\frac{1}{2}V_0$$
$$C = 0 \qquad \gamma = D.$$

Finally she uses conditions (1) and (2) to find the feasible set, then chooses the smallest allowable x to be the optimal solution (the solution with least fuel used). The working used is as follows (keeping in mind that all the constants are positive):

Condition (1) implies: $\frac{1}{2}(Fx - G) > 0$, so

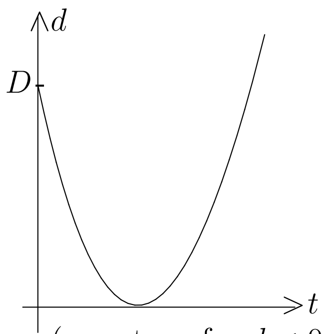
$$x > \frac{G}{F}$$

Condition (2) implies: $\frac{1}{2}D(Fx - G) - (-V_0)^2 \geq 0$, so

$$x \geq \frac{GD + 2V_0^2}{FD}.$$

Since both conditions (1) and (2) must hold, the larger of the two is chosen. The latter of the two is always larger considering $\frac{GD+2V_0^2}{FD} = \frac{G}{F} + \frac{2V_0^2}{FD}$. The pilot concludes that the rate that fuel should be used is $x^* = \frac{GD+2V_0^2}{FD}$. The feasible set in this case is not bounded (feasible x can go to positive infinity), but because it is bounded at the lower end the optimal solution exists. It is easy to verify that this course of action will bring her to just touch the point of "no-return".

Optimal Path



(no return for $d < 0$)

In this description of semi-definite optimization, only problems of one solution variable have been considered. It is for this reason that the example described above is of little practical significance. Semi-definite optimization can be extended to an unlimited number of variables and is usually quite difficult to solve. However, the format of this problem is commonly arrived at in practical applications; these may have hundreds and even thousands of variables. One such application is the design of optimal building trusses (such as those commonly used in roof structures). In this application the problem is to find the structure that best supports a set of loads, with configuration constraints such as a limit on the total amount of material that can be used.

References

[1]R. Fletcher, *Practical Methods of Optimization*, John Wiley and Sons, 1991.

[2] L. Vandenberghe and S. Boyd, *Semi-definite Programming*, SIAM Review 38 (1996) 49-95.