Lecture 7: Maximising $c^T x$ Over a Convex Set

Lecturer: *Sundar Vishwanathan*
Computer Science & Engineering Indian Institute of Technology, Bombay

1 Convex Sets

Recall the definitions of a convex set from the previous lecture.

Theorem 1 *If S_1 and S_2 are two convex sets, then $S_1 \cap S_2$ is a convex set.*

Proof: Let $x_1, x_2 \in S_1 \cap S_2$. Now since x_1 and x_2 belong to S_1 (which is convex), any convex combination of them lies in S_1. Similarly we can say that this convex combination of x_1 and x_2 lies in S_2. Thus the convex combination lies in $S_1 \cap S_2$. Thus $S_1 \cap S_2$ is convex. □

We can use the previous theorem to give a proof of the fact that $\{x : Ax \leq b\}$ is a convex set. It is, in a sense, the same as the previous proof but easier to visualise. Consider $A_1 x \leq b_1$. All the points satisfying this inequality lie on one side of the hyperplane $A_1 x = b_1$. This set is convex. If we take two points satisfying this inequality, it can be easily checked that so does every convex combination. Similarly the solution sets of the other inequalities $A_j x \leq b_j$ are also convex. Thus, $Ax \leq b$, which is an intersection of all of these convex regions, is convex.

2 Maximize $c^T x$

Recall our quest. How do we maximize $c^T x$ over the set of all x satisfying $Ax \leq b$? We now know that $Ax \leq b$ is a convex set. In subsequent lectures we will have more to say about what it looks like. It is instructive to see how the value of the function $c^T x$ varies in R^n. This is something you know implicitly but we wish to make it explicit.

For creating a picture in your mind, we will initially restrict attention to two dimensions. So consider $c_1 x + c_2 y$ as it varies over R^2. We know that $c_1 x + c_2 y = 0$ is a line through the origin. And for any value α, $c_1 x + c_2 y = \alpha$ is a line which is parallel to the above line. So, the value of this function is zero on the line $c_1 x + c_2 y = 0$, increases monotonically in a direction perpendicular to the line and decreases in the reverse direction.

Indeed, this picture is true even in higher dimensions. Can we identify a vector in the direction perpendicular to $c^T x = \beta$? A vector perpendicular to $c^T x = \beta$ is also perpendicular to $c^T x = 0$. Why is this? Can you now identify a vector perpendicular to $c^T x = 0$?

What does the hyperplane $c^T x = 0$ mean? It is the collection of all points which are perpendicular to the vector c. Restating this, the vector c is perpendicular to all points on the hyperplane $c^T x = 0$. And this is the vector we are looking for. In fact, it is easy to see that the value increases along the direction of c and decreases along $-c$.

Maximising this function on a convex set is not difficult to imagine. Let us take a simple convex set: the unit sphere given by $x^T x \leq 1$.

This is a set such that all points are at a distance less than or equal to 1 from the origin.
Exercise: Prove that this set is convex.
We know that $c^T x$ increases in the direction of c. Thus we start moving in the direction of c. The last point where $c^T x$ touches the sphere is the point of maxima. It can be easily seen that at this point, $c^T x$ is a tangent to the sphere $x^T x$.

Now consider any convex polygon in the plane. To maximize $c^T x$ keep moving along c. The last point where $c^T x$ touches the polygon will be the point of maxima. It can be observed that this point will be a point on the boundary.
While we have made several loose statements (for instance what is the definition of boundary ?) the picture we have painted is kosher and we will make all this formal in the forthcoming lectures.
Generously extending this argument to the case of our interest, it is natural to think that $c^T x$ will attain its maximum value at some boundary point(s) of the region $Ax \leq b$. This is indeed true and we will prove this too. In fact we will see shortly that $Ax \leq b$ in 2D looks like a polygon.

A function $f : R^n \to R$ is called *linear* if it satisfies the following:

1. $f(x + y) = f(x) + f(y)$ and
2. for any real α, $f(\alpha x) = \alpha f(x)$.

Exercise: Check that the function $g_c : R^n \to R$ defined by $g_c(x) = c^T x$ is linear.

A possibly surprising fact is that all linear functions are of the form g_c for some c. Why is this? Reverse engineering is called for to prove this. If indeed this were true, what would c be? Indeed, what should the first coordinate of c be in terms of f?
You need to note now that if this were true then c_i must be $f(e_i)$.

Solve the exercise below to finish the proof. *Exercise*: Prove that $f(x) = (f(e_1), f(e_2), \ldots, f(e_n))^T x$, for all x.

Maximising an arbitrary function over an arbitrary infinite set is an intractable computational problem. However, in our case, both the function and the sets are well behaved. We illustrate one key property below.

DEFINITION 1 *A point x_0 in a set S is said to be a local maxima for a function f if there exists a small neighbourhood N of x_0 where $f(x_0) \geq f(x), \forall x \in N$. For us, N is a ball of a small but non-zero radius around x_0.*

THEOREM 2 *Let f be a linear function over a convex set S. Then a local maximum is a global maximum.*

PROOF: Let x_0 be a local maximum and y be a global maximum. The basic idea is to consider the line segment between x and y and show that the value of the function varies continuously in a non-decreasing way from x_0 to y. This is a property of linear functions. The existence of the line segment inside the set is due to the convexity property.
Let N be a neighbourhood where x_0 is the local maximum. Consider a point $P = (1-\epsilon)x_0 + \epsilon y$. Choose ϵ small enough so that this point lies inside N.
Now, consider $f((1-\epsilon)x_0 + \epsilon y)$. Since f is linear, this can be written as

$$(1-\epsilon)f(x_0) + \epsilon f(y)$$

$$= f(x_0) + \epsilon(f(y) - f(x_0))$$

Observe that at point P, which is in the neighbourhood of x_0, the value of the function is strictly greater than $f(x_0)$ if $f(y) > f(x_0)$. Thus $f(x_0)$ can be maximum only if $f(x_0) = f(y)$. Or that a local maximum is the same as a global maximum. $\square$