1. Introduction to convex optimization

Historically, linear programs were the focus in optimizaiton.
The dividing line betwen tractable and intractable problems was considered to be linear vs non-linear.
But many non-linear problems turned out to be quite tractable.
Increasingly the right distinction is seen as being between convex vs non-convex problems.

2. Convex sets and functions

A convex set $C \subseteq \Re^n$ has the following property.
Take any two points in the set: x, y. Any line segment between the two points has to reside entirely in the set.
That is, $tx + (1-t)y \in C \; for\; all\; 0 \leq t \leq 1$

Everything about convex functions can be dervied from a property of convex sets.
A convex function $f : \Re^n \to \Re$ is a function such that $domain(f) \subseteq \Re^n$ is convex.
And if you take two points in the domain, the function value
$f(tx + (1-t)y) \leq tf(x) + (1-t)f(y)$ for all $0 \leq t \leq 1$.
i.e. the function evaluated at every point between two points in the domain must always be less than the line segment joining the two points.

What is the domain of a function?

- set of all x such that $f(x)$ is defined and finite.
- sometimes we may drop this: e.g. for $f(x) = \log x$, its domain is $\Re^{++}$i.e. all $x \geq 0$.

3. Generic form of optimization problem
$min_{x \in D} \; f(x)$

$subject\; to$
$g_i(x) \leq 0, \; i = 1, \ldots, m$
$and\; h_j(x) = 0, \; j = 1, \ldots, r$

This is a **convex optimization** problem if:
1. the functions f and g_i are convex
2. h_j are affine, i.e. $h_j(x) = a_j^T x + b_j$

In every convex opt problem: **local minima are global minima**. This is what makes them tractable.

Theorem: If x is feasible ($x \in D$ and it satisfies all constraints), and it minimizes f in a local neighbourhood, i.e. $f(x) \leq f(y)$ *for all feasible* y, $||x - y||_2 \leq \rho$
then we know that $f(x) \leq f(y)$ for all feasible y.

Proof: by contradiction.
Suppose there is some feasible point z st $f(z) < f(x)$. (strictly smaller)
Then we know $||x - z||_2 \geq \rho$, since x is locally optimal within the radius ρ around x.
Consider $tx + (1-t)z = y$. These are points on the line segment between x and z.
First we will show that y is feasible.
Fact 1: $y \in D$. This is because D is an intersection of convex sets, which are also convex.
Fact 2: $g_i(y) \leq 0$. This is because $g_i(tx + (1-t)z) \leq tg_i(x) + (1-t)g(z) \leq 0$, since by definition g_i is a convex function, and these are the conditions in the original model spec.
(use convexity of g)
Fact 3: $h_j(y) = 0$. Why? (use affine/convexity of h)
Hence, y is feasible, that is, everything on the line segment joining x to z is feasible for the optimization problem.
We know that
$||y - x||_2 = ||tx + (1-t)z - x||_2$
$= (1-t)\ ||z - x||_2$
Therefore, we can choose t close to 1 st $||y - x||_2 \leq \rho$. That is, we can get y within the ball around x of radius ρ.

We know f is convex. (use convexity of f) So
$f(y) = f(tx + (1-t)z) \leq tf(x) + (1-t)f(z) < tf(x) + (1-t)f(x) = f(x)$, because we assume that $f(z) < f(x)$.

- This property does not exist for non-convext problems! Which makes things trickier.

But then we have $f(y) < f(x)$ and $||y - x||_2 < \rho$, which contradicts the earlier statement that x is locally optimal in the ball. Hence no such point z can exist.