1 30th of November 2018 — A. Frangioni

1.1 Constrained optimization

In this lecture we address the problem of finding the **optimum** of a function in a subset of its domain, called X. The term optimum differs from the minimum, because the optimum in that subset may not be a minimum of the whole function.

$f_* = \min\{ f(x) \ : \ x \in X \}$

Definition 1.1 (Local optimum)**.** *Given a function f and a constraint set X, we denote* ***local optimum*** *the point where the function assumes the minimum value inside the set X. Formally,* $\min\{f(x) \ : \ x \in \mathcal{B}(x_*, \varepsilon) \ \cap \ X\}$ *for some* $\varepsilon > 0$.

Notice that the only points in which the constraint adds some informations are the ones on the boundary, as shown in Figure 1.1

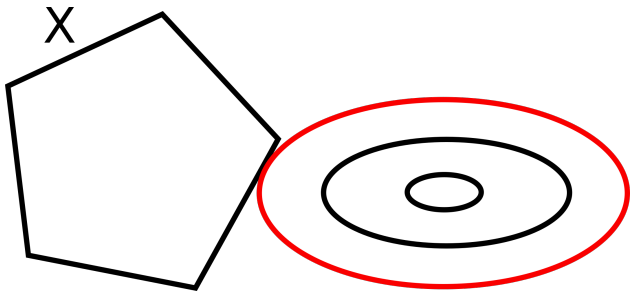

FIGURE 1.1: The red line is level set of the function corresponding to the smallest value that touches the set X. The point in the intersection is not a saddle point of the function f, altough it is the minimum.

There are two kinds of contraints:

FAKE ONES: this first kind is such that the minimum of the function lies inside the set X, hence there is no need to use the constraints at all;

REAL ONES: when the optimal is on the boundary. This is the case of linear functions, because the gradient is constant $\nabla f(x) = c$.

At this point we want to decide if a point on the boundary is an optimum. In this context it is important how the boundary is defined.

1.1.1 Linear equality constraints

A constraint of this kind is very simple: it is a subspace, as shown in Figure 1.2.

$\min\{f(x) \ : \ Ax = b\}$, where the rank of A counts the number of linearly independent rows.

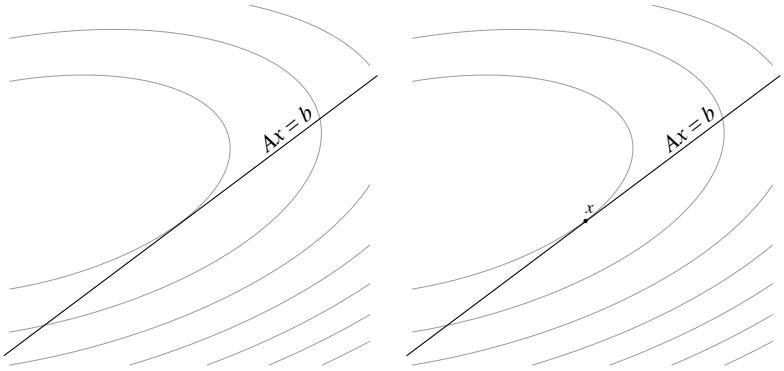

FIGURE 1.2: Linear constraint and a point on the boundary.

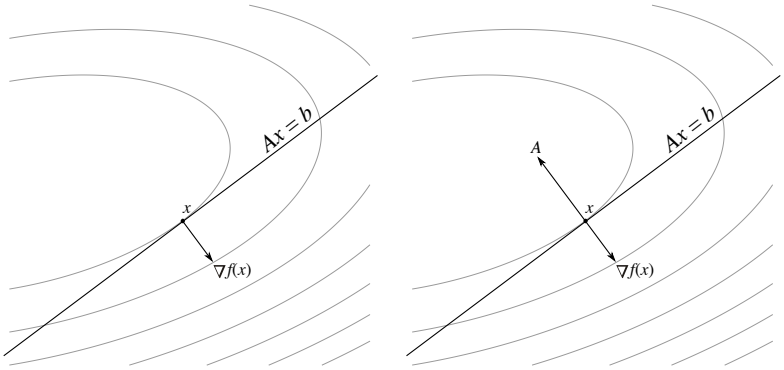

FIGURE 1.3: The gradient is orthogonal to the level set in that point, when the function is smooth. The same holds for matrix A.

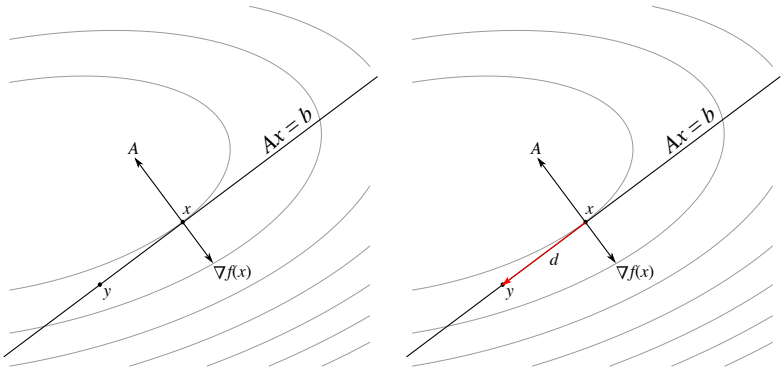

FIGURE 1.4: If we take any other point in the space it has to be orthogonal to A.

Let us assume that there are not linearly independent rows in A, then or this behaviour is reflected in B or the system does not have any solution.

In the case of presence of linearly dependent columns such columns may be eliminated to ease the computation without loss of information.

The intuition behind what follows is that each linear constraint kills one degree of freedom, formally $det(A_B) \neq 0 \Rightarrow Ax = b \equiv x_B = A_B^{-1}(b - A_N x_N) \Rightarrow$.

We want to extract a submatrix A_B from $A \in M(m,n,\mathbb{R})$, such that $A_B \in M(m,\mathbb{R})$ and then the system induces a partitioning in the variables as well.

$A = [A_B, A_N]$, $x = [x_B, x_N]$, so the system becomes $A_B x_B + A_N x_N = b$ and (since A_B is non singular) $X_B + {A_B}^{-1} x_N = {A_B}^{-1} b$ in other words, given the indepedent variables we can compute the values of the dependent ones and this is a linear operation.

The original optimization problem becomes an optimization problem on a reduced space, formally $\min\{r(w) = f(Dw+d) : w \in \mathbb{R}^{n-m}\}$, where $D = \begin{bmatrix} -A_B^{-1}A_N \\ I \end{bmatrix}$ and $d = \begin{bmatrix} A_B^{-1}b \\ 0 \end{bmatrix}$. For each point in the smaller space we can compute the function in the larger space.

How can we compute the gradient of $r(w)$? The gradient is $\nabla\, r(w) = D^T \nabla\, f(Dw+d)$. The fact that w^* is an optimum implies that $\nabla\, r(w^*) = 0$.

$D = \begin{bmatrix} -A_B^{-1}A_N \\ I \end{bmatrix}$ then $AD = [A_B, A_N] \cdot \begin{bmatrix} -A_B^{-1}A_N \\ I \end{bmatrix} = -A_B \cdot {A_B}^{-1} \cdot A_N + A_N = 0.$

Now the point is taking a multiple of matrix A, finding a feasible x and corresponding w (because there is a bijection), then finding the value μ that allows the equality.

Theorem 1.1. *Let $Ax = b$ be a linear system and w such that $x = \left[{A_B}^{-1}(b - A_N w), w\right]$. If $\exists\ \mu\ \in\ \mathbb{R}^m$ s.t. $\mu A = \nabla\ f(x)$ then $r(w) = D^T \nabla\ f(x) = 0$, see Figure 1.5. In other words, it is equivalent to find a stationary point x for the original problem (P) or finding the stationary point w for (R).*

Definition 1.2 (Poorman's Karush Kuntaker conditions)**.** *A point is a good candidate for being a minimum of the constrained problem if and only if it satisfy **Poorman's KKT conditions**, namely the problem is feasible and that $\exists \mu \in \mathbb{R}^m$ s.t. $\mu A = \nabla f(x)$.*

Theorem 1.2. *Let f be a convex function, then KKT conditions are enough for optimality.*

A very naive explanation of the theorem is that if the function is convex also the restriction is convex and a tationary point of a convex function is a minimum.

Our idea is to characterize the directions we can move along in order to find new points that satisy the constraint. Formally, $Ax = b$ is our contstraint and we want to move towards $x + d$ and stay in the feasible region. How? $A(w+d) = b \Leftrightarrow Ax + Ad = b \Leftrightarrow Ad = 0$, since $Ax = b$. The only way to move along the constraint is choosing a direction which scalar product with A is 0, hence 0 scalar product with the gradient.

From now on we would like to study the behaviour on contrained problems where the constraints are equalities, but inequalities.

In order to do this we need some mathematical background.

1.1.2 Background for linear inequality contstraints

Definition 1.3 (Tangent cone)**.** *We call **tangent cone** of X at x $\mathbf{T_X(x)}$ =t.*

$$\{d \in \mathbb{R}^n\ :\ \exists\ \{z_i \in X\} \to x\ \wedge\ \{t_i \geq 0\} \to 0\ s.t.\ d = \lim_{i\to\infty} \frac{z_i - x}{t_i}\}$$

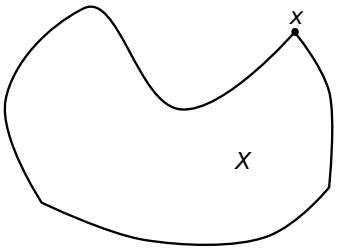

FIGURE 1.5: When the boundary has this shape we can move along directions that point inside the constraints. Tangent directions are not allowed.

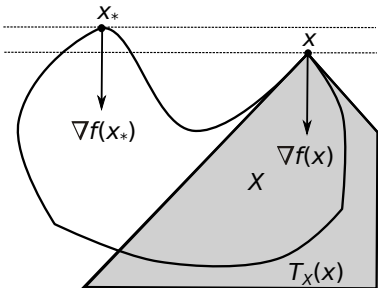

FIGURE 1.6: Geometric representation of tangent cone, which is the region of the space where we can pick the directions. The intuition is to zoom in x and the result of this zooming is a cone.

Theorem 1.3. *Let $\mathcal{C}$ be a cone, $\forall x \in \mathcal{C}\ \alpha x \in \mathcal{C},\ \forall \alpha > 0$.*

Theorem 1.4. *Given a function f, where x is a **local** optimum $< \nabla f(x), d > \geq 0\ \forall d \in T_X(x)$.*

Proof. **Proof by contraddiction:** Assume $\exists d \in T_X(x)$ such that $< \nabla f(x), d >< 0$, but x is a local optimum.

By definition $\exists X \supset \{z_i\} \to x$ and $\{t_i\} \to 0$ such that $d \overset{*}{=} \lim_{i\to\infty} \frac{z_i - x}{t_i}$.
First order Taylor $f(z_i) - f(x) =< \nabla f(x), (z_i - x) > + R(z_i - x)$.

$$\begin{aligned}
\lim_{i\to\infty} \frac{f(z_i - x)}{t_i} &= \lim_{i\to\infty} < \nabla f(x), \frac{z_i - x}{t_i} > + \frac{R(z_i - x)}{t_i} \\
&\overset{*}{=} < \nabla f(x), d > + \lim_{i\to\infty} \frac{R(z_i - x)}{t_i} \\
&\overset{(1)}{=} < \nabla f(x), d > \\
&< 0
\end{aligned} \tag{1.1}$$

Where, $\overset{(1)}{=}$ follows from $\lim_{i\to\infty} \frac{R(z_i - x)}{t_i} = 0$ by Taylor. □

Observation 1.1. *The optimum of Theorem 1.4 is global when the function is convex, because in that case $X \subseteq x + T_X(x)$. For a geometric idea see Figure 1.7.*

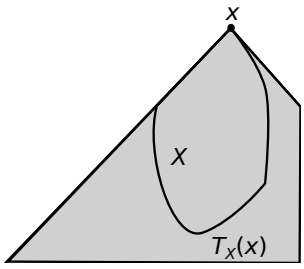

FIGURE 1.7: Convex function.

Observation 1.2. *Notice that the rule* $< \nabla f(x), d > \geq 0 \ \forall d \in T_X(x) \Rightarrow x$ *local optimum does not hold. Let us see a counter example:* $\min\{ x_2 : x_2 \geq x_1^3 \}$*, displayed in Figure 1.9.*

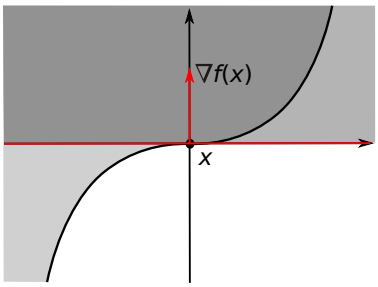

FIGURE 1.8: Let us suppose we pick the direction towards the left part of the function in the saddle point. That direction is promising and actually the value of the function decreases. In this case the problem is that the contraint is not convex.

Now we need a more manageable object for $T_X(x)$.

Definition 1.4 (Cone of feasible directions)**.**
Intuitively, we are in x *and we want to find al directions* ***feasible cone*** *such that there exist small but not* 0 *steps such that all the points on this direction are feasible. Formally, a* ***feasible cone*** *is* $F_X(x) = \{d \in \mathbb{R}^n \ : \ \exists \bar{\varepsilon} > 0$ *such that* $x + \varepsilon d \in X, \ \forall \varepsilon \in [0, \bar{\varepsilon}]\}$*.*

Fact 1.5. *The properties of such cone are:*

1. T_X *closed,* F_X *in general not (hence the cone of feasible directions is the tangent cone minus the tangent directions);*
2. $cl(F_X) \subseteq T_X$*, where* $cl(F_X)$ *is the closure of the cone of feasible directions;*
3. *If* X *convex then the cones coincide:* T_X *and* F_X *convex and* $cl\, F_X = T_X$*.*

We are now interested in finding a better characterization of the cone fo feasible directions, hence we introduce a new characterization fo the set of constraints X.

First representation:

$$X = \{x \in \mathbb{R}^n \ : \ g_i(x) \leq 0 \ \ i \in \mathcal{I}, \ h_j(x) = 0 \ \ j \in \mathcal{J} \}$$

where $\mathcal{I}$ is the set of inequality constraints and $\mathcal{J}$ is the set of equality constraints;

Second representation:

$$X = \{ \, x \in \mathbb{R}^n \ : \ G(x) \leq 0 \, , \, H(x) = 0\}$$

where $G = [g_i(x)]_{i \in \mathcal{I}} : \mathbb{R}^n \to \mathbb{R}^{|\mathcal{I}|}$ and $H = [h_i(x)]_{i \in \mathcal{J}} : \mathbb{R}^n \to \mathbb{R}^{|\mathcal{J}|}$;

Third representation: (hiding equalities)

$$X = \{x \in \mathbb{R}^n : g_i(x) \leq 0 \ i \in \mathcal{I}, \ h_j(x) \leq 0 \ \wedge \ h_j(x) \geq 0 \ j \in \mathcal{J}\}$$

Fourth representation: (hiding inequalities into a single function)

$$X = \{x \in \mathbb{R}^n \ : \ g(x) = \max\{ \, g_i(x) \ : \ i \in \mathcal{I} \, \} \leq 0 \ i \in \mathcal{I}, \ h_j(x) = 0 \ \in \mathcal{J}\}$$

Definition 1.5 (Active constraints)**.** *We term* ***active constraints*** *at $x \in X$ the following set*

$$\mathcal{A}(x) = \{i \in \mathcal{I} \ : \ g_i(x) = 0\} \subseteq \mathcal{I}$$

Let us introduce some useful notation on the subject: let $\mathcal{B} \subseteq \mathcal{I}$ a subset of indices. We denote $G_{\mathcal{B}} = [g_i(x)]_{i \in \mathcal{B}} : \mathbb{R}^n \to \mathbb{R}^{|\mathcal{B}|}$ the corresponding set of inequalities.

Definition 1.6 (First-order feasible direction cone)**.** *We term* ***First-order feasible direction cone at*** *$x \in X$:*

$$D_X(x) = \{d \in \mathbb{R}^n \ : < \nabla g_i(x), d > \leq 0 \ i \in \mathcal{A}(x) \ < \nabla h_j(x), d > = 0 \ j \in \mathcal{J}\} = \{d \in \mathbb{R}^n : (JG_{\mathcal{A}(x)}(x))d \leq 0,$$

Intuitively, the fact that we are looking at the active set means that we are zooming very close to 0.

In this cone we require that all the directions inside it have a negative scalar product with the gradient of the contraints.

A visual example of a first order feasible direction cone is displayed in Figure 1.9.

Fact 1.6. *The tangent cone is a subset of the first-order feasible direction cone. Formally, $T_X(x) \subseteq D_X(x)$.*

We would like the first-order feasible direction cone to be exactly equal to the tangent cone and this is almost always true, except for some pathological cases.

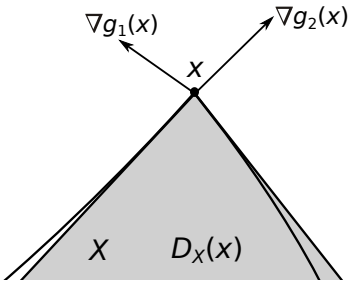

FIGURE 1.9: The first-order feasible direction cone is made of those directions that are orthogonal to the gradient of the constraints $g_1(x)$ and $g_2(x)$ in x.

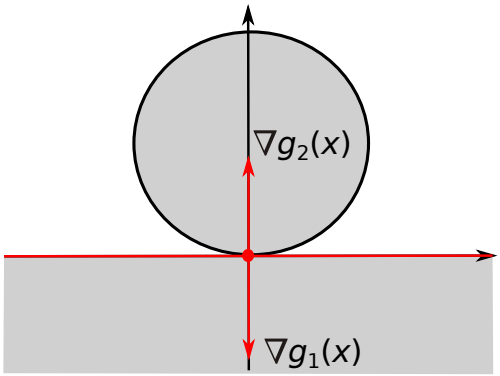

FIGURE 1.10: The circle represents the quadrati constraint, while the semi-plane represents the first degree constraint.

Example 1.1. *Let our minimum problem be* $\min\{\dots : {x_1}^2 + (x_2 - 1)^2 - 1 \leq 0,\ x_2 \leq 0\}$.
The plot of such functions is shown in Figure 1.10.
Our claim is that the only feasible point is $X = \{x = [0, 0]\}$.
The only feasible direction is 0*, hence cone of feasible direction and the tangent cone are the singleton* $\{[0, 0]\}$.
On the other hand, the set of directions that have non-negative scalar product with both g_1 *and* g_2 *are all the* x *axis.*

We would like to ensure we are not in one of these pathological cases and to do this we introduce some conditions.

Fact 1.7. *The following holds:*

AFFINE CONSTRAINTS (AFFC): *Let* g_i *and* h_j *be affine constraints. Then,* $\forall\ i \in \mathcal{I}$ *and* $j \in \mathcal{J}$ $T_X(x) = D_X(x)\ \forall x \in X$.

SLATER'S CONDITION (SLAC): *Let* g_i *convex* $\forall\ i \in \mathcal{I}$ *and let* h_j *affine* $\forall\ j \in \mathcal{J}\ \exists\ \bar{x} \in X$ *s.t.* $g_i(\bar{x}) < 0\ \forall\ i \in \mathcal{I}$. *Then* $T_X(x) = D_X(x)\ \forall\, x \in X$;

LINEAR INDEPENDENCE (LINI): $\bar{x} \in X \wedge$ *the vectors* $\{\nabla g_i(\bar{x}) \ : \ i \in \mathcal{A}(\bar{x})\} \cup \{\nabla h_j(\bar{x}) \ : \ j \in \mathcal{J}\}$ *linearly independent* $\Longrightarrow T_X(\bar{x}) = D_X(\bar{x})$. *Among all these conditions this is the only local one.*

It goes without saying that we cannot check all the directions in order to exclude the nasty pathological cases.

Definition 1.7 (Dual cone)**.** *Let* D_X *be a* ***polyhedral cone*** $\mathcal{C} = \{d \in \mathbb{R}^n \ : \ Ad \leq 0\}$, *for some* $A \in \mathbb{R}^{k \times n}$.

We term ***dual cone*** $\mathcal{C}^* = \{c = \sum_{i=1}^{k} \lambda_i A_i \ : \ \lambda \geq 0 \}$.

Lemma 1.8 (Farka's lemma)**.** *Intuitively, this lemma says that pick a vector: either it belongs to the dual cone or there exists a vector in the polyhedral cone which has a negative scalar product with it.*

Equivalently, either $c \in \mathcal{C}^*$ *or* $c \notin \mathcal{C}^*$.

More formally, either $\exists\, \lambda \geq 0$ *s.t.* $c = \sum_{i=1}^{k} \lambda_i A_i$ *or* $\exists\, d$ *s.t.* $Ad \leq 0 \wedge < c, d >> 0$.

Theorem 1.9 (Karush-Kuhn-Tucker conditions)**.** *Let us assume that we found an optimal solution* x_* *and the constraints qualification holds.*

Then $\exists \lambda \in \mathbb{R}_+^{|\mathcal{I}|}$ *and* $\mu \in \mathbb{R}^{|\mathcal{I}|}$ *such that*

$$\nabla f(x) + \sum_{i \in \mathcal{A}(x)} \lambda_i \nabla g_i(x) + \sum_{j \in \mathcal{J}} \mu_j \nabla h_j(x) = 0$$

It is interesting to notice that we did not impose $\mu \geq 0$. Let us take an equality constraint $h_j(x) = 0$. This is equivalent to write $h_j(x) \leq 0 \ \wedge h_j(x) \geq 0$, thus leading to two different multipliers, say λ_j^+ and λ_j^-.

The term of the sum concerning h_j looks like this $\lambda_j^+ \nabla h_j(x_*) - \lambda_j^- \nabla h_j(x_*) = (\lambda_j^+ - \lambda_j^-) \cdot \nabla h_j(x_*)$, where both λ_j^+ and λ_j^- are ≥ 0, hence their difference (denoted by μ_j) may be either positive or negative.

Fact 1.10. *The Karush-Kuhn-Tucker conditions are also written as:*

FEASIBILITY: $x \in X \equiv g_i(x) \leq 0 \ \ i \in \mathcal{I}, \ h_j(x) = 0 \ \ j \in \mathcal{J}$

KKT-G: $\nabla f(x) + \sum\limits_{i \in \mathcal{I}} \lambda_i \nabla g_i(x) + \sum\limits_{j \in \mathcal{J}} \mu_j \nabla h_j(x) = 0$

COMPLEMENTARITY SLACKNESS: $\sum\limits_{i \in \mathcal{I}} \lambda_i g_i(x) = 0$

where the second and the third equations formalize the definition of KKT, where the terms of the first sum should be summed only if their contraints are active.

Fact 1.11. *Let (P) be a convex problem. In this case, if the Karush-Kuhn-Tucker conditions hold then* x *is a global optimum.*