Convexity, and Hung's inequality with linear constraints.

January 24^{th} 2017

Martin Celli
Departamento de Matemáticas
Universidad Autónoma Metropolitana-Iztapalapa
Av. San Rafael Atlixco, 186. Col. Vicentina.
CP 09340. Del. Iztapalapa, Ciudad de México.
E-mail: cell@xanum.uam.mx

The aim of this article is to state and demonstrate a theorem on the optimization of convex functions, which can be useful to any lector keen on elementary inequalities. It will give us two proofs of an example of an inequality with linear constraints: one of them using calculus, the other using only algebra. This article is also an invitation to enjoy the geometric structures hidden behind any inequality.

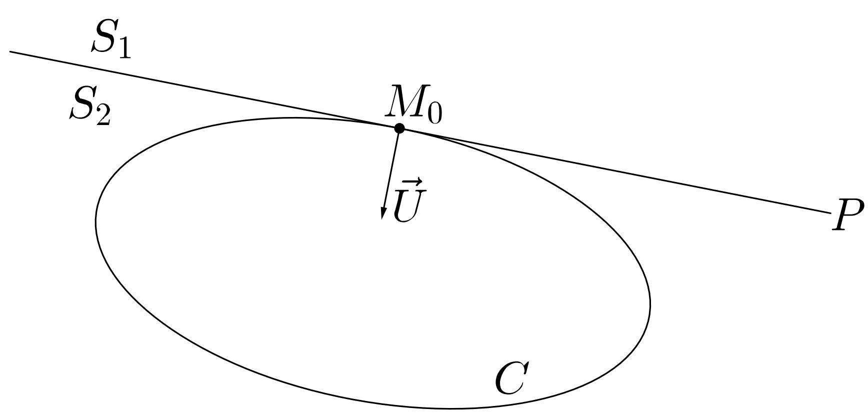

Figure 1

Let us start with the following result:

Theorem 1. *Let Ω be an open subset of $\mathbb{R}^N$, F a differentiable function: $\Omega \to \mathbb{R}$, C a convex subset of Ω (for every A, $B \in C$, the line segment $[A, B]$ is a subset of C), M_0 a point of C, $\vec{U} = \vec{\nabla}_{M_0} F$ the gradient of F at the point M_0. Let us assume that $\vec{U} \neq \vec{0}$. Let P be the hyperplane passing through M_0 and having $\vec{U}$ as a normal vector, S_1 and S_2 the half-spaces defined by P (such that $\vec{U}$ is oriented from S_1 to S_2). Then we have the following implication:*
if the restriction of F to C achieves its minimum value at M_0, then $C \subset S_2 \cup P$ (figure 1).

In order to prove this theorem, let us consider a possible point M of $C \cap S_1$ (figure 2). Let $\vec{V} = \overrightarrow{M_0M}$. By the definition of the gradient, we have, for all $t \approx 0$:

$$F(M_0 + t\vec{V}) = F(M_0) + t(\vec{U}|\vec{V}) + ||t\vec{V}||\epsilon(t\vec{V}),$$

where $(.|.)$ denotes the standard scalar product of $\mathbb{R}^N$ and ϵ is a function whose limit at $\vec{0}$ is 0.

$$F(M_0 + t\vec{V}) - F(M_0) = t\left((\vec{U}|\vec{V}) + \left(\frac{|t|}{t}\right)||\vec{V}||\epsilon(t\vec{V})\right) \sim t(\vec{U}|\vec{V})$$

when $t \to 0$, for $(\vec{U}|\vec{V}) \neq 0$. As $M_0 \in S_1$: $(\vec{U}|\vec{V}) < 0$. Consequently, for $t \gtrsim 0$, we have:

$$F(M_0 + t\vec{V}) - F(M_0) < 0,$$

$$F(M_0 + t\vec{V}) < F(M_0)\cdot$$

As C is convex and $0 < t < 1$, the point $M_0 + t\vec{V}$ belongs to C. Thus, the restriction of F to C cannot achieve its minimum value at M_0. ■

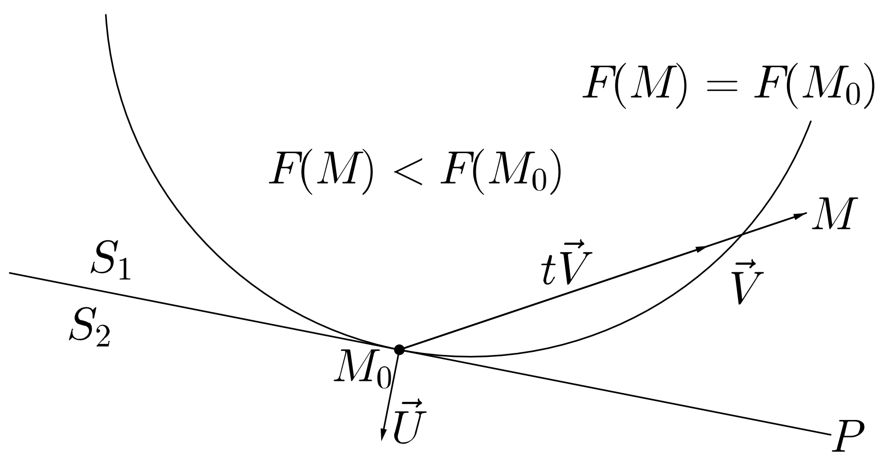

Figure 2

Unfortunately, the converse of this theorem is false, unless we add a hypothesis on F. Let us define the Hessian matrix of a function $F(x_1, \ldots, x_N)$ as the matrix:

$$H_{(x_1,\ldots,x_N)}F = \left(\frac{\partial^2 F}{\partial x_i \partial x_j}\right)_{1 \leq i,j \leq N},$$

when defined. It can be shown that, if F is twice differentiable:

$$\frac{\partial^2 F}{\partial x_i \partial x_j} = \frac{\partial^2 F}{\partial x_j \partial x_i}\cdot$$

This result is proved in [1], in the particular case where F is a $\mathcal{C}^2$ function. It implies that $H_{(x_1,\ldots,x_N)}F$ is symmetric. So $H_{(x_1,\ldots,x_N)}F$ is orthogonally diagonalizable for the standard scalar product of $\mathbb{R}^N$ (for a proof, see [2]). Let us define a convex function as a twice differentiable function whose Hessian matrix has only positive (≥ 0) eigenvalues at every point M_0. As $H_{M_0}F$ is symmetric, it is equivalent to say that $(H_{M_0}F\vec{V}|\vec{V}) \geq 0$ for every vector $\vec{V}$.

The following result is a converse of Theorem 1, under a hypothesis of convexity.

Theorem 2. *Under the hypotheses of Theorem 1, we have the following implication: if F is convex and $\mathcal{C}^2$ in C and $C \subset S_2 \cup P$, then the restriction of F to C achieves its minimum value at M_0.*

In order to prove this theorem, as F is $\mathcal{C}^2$ in C, we can apply the Taylor formula (proved in [1]) to F at order 2 at the point M_0. Thus, for every $M \in C$, there is a point M' in the open line segment (M_0, M) (which is a subset of C as C is convex) such that:

$$F(M) = F(M_0) + (\vec{U}|\overrightarrow{M_0M}) + \frac{1}{2}(H_{M'}F\overrightarrow{M_0M}|\overrightarrow{M_0M})\cdot$$

As F is convex, the quadratic term is ≥ 0:

$$F(M) \geq F(M_0) + (\vec{U}|\overrightarrow{M_0M})\cdot$$

In $\mathbb{R}$ and $\mathbb{R}^2$, this is equivalent to say that the graph of a convex function always lies above its tangent lines and planes. Now we just have to see that the linear term is ≥ 0 in $S_2 \cup P$. ■

We can use this theorem to prove various inequalities, such as the following, due to Nguyen Viet Hung:

Proposition. *Let a, b, c be three positive (> 0) real numbers such that:*

$$\begin{cases} a \leq 2 \\ a+b \leq 5 \\ a+b+c \leq 11 \end{cases}\cdot$$

Then we have:

$$\frac{1}{a} + \frac{1}{b} + \frac{1}{c} \geq 1\cdot$$

Several proofs of this result can be found in [3]. We can also apply Theorem 2 to:

$$\Omega = (0, +\infty)^3, F(a,b,c) = \frac{1}{a} + \frac{1}{b} + \frac{1}{c},$$

$$C = \left\{ (a,b,c) \in (0,+\infty)^3, \begin{cases} a \leq 2 \\ a+b \leq 5 \\ a+b+c \leq 11 \end{cases} \right\}, M_0 = (2,3,6)\cdot$$

Then $\vec{U} = (-1/4, -1/9, -1/36)$ and S_2 is the half-space of points $M = (a,b,c)$ such that:

$$0 < \vec{U}.\overrightarrow{M_0M} = \frac{2-a}{4} + \frac{3-b}{9} + \frac{6-c}{36}\cdot$$

This is equivalent to: $9a+4b+c<36$. Thus we have:

$$S_2 \cup P = \{(a,b,c) \in \mathbb{R}^3, 9a+4b+c \leq 36\}\cdot$$

In order to show that $C \subset S_2 \cup P$, we just have to see that, in C:

$$9a+4b+c = 5a+3(a+b)+(a+b+c) \leq 5\times 2+3\times 5+11 = 36\cdot \blacksquare$$

We are now going to convert this proof to a two-line proof which does not involve calculus. In order to do so, we have to compute three second-order Taylor remainders:

$$\frac{1}{a}-\frac{1}{2}-\left(\frac{1}{a}\right)'_{a=2}(a-2) = \frac{1}{a}-\frac{1}{2}+\frac{a-2}{2^2} = \frac{(2-a)^2}{2^2a}\cdot$$

Thus:

$$\frac{1}{a} = \frac{1}{2}+\frac{2-a}{4}+\frac{(2-a)^2}{4a}\cdot$$

Similarly, we have:

$$\frac{1}{b} = \frac{1}{3}+\frac{3-b}{9}+\frac{(3-b)^2}{9b},$$

$$\frac{1}{c} = \frac{1}{6}+\frac{6-c}{36}+\frac{(6-c)^2}{36c}\cdot$$

Here is our elementary two-line proof:

$$\frac{1}{a}+\frac{1}{b}+\frac{1}{c} = \left(\frac{1}{2}+\frac{2-a}{4}+\frac{(2-a)^2}{4a}\right)+\left(\frac{1}{3}+\frac{3-b}{9}+\frac{(3-b)^2}{9b}\right)+\left(\frac{1}{6}+\frac{6-c}{36}+\frac{(6-c)^2}{36c}\right)$$

$$=1+\frac{5(2-a)+3(5-(a+b))+(11-(a+b+c))}{36}+\frac{(2-a)^2}{4a}+\frac{(3-b)^2}{9b}+\frac{(6-c)^2}{36c} \geq 1,$$

as the last four terms are ≥ 0.

References.

[1] C. H. Edwards Jr., Advanced Calculus of Several Variables. Dover Publications, 1994.
[2] S. Lang, Linear Algebra. Springer, Undergraduate Texts in Mathematics, 2004.
[3] A. Bogomolny, “Inequality with Three Linear Constraints”. Interactive Mathematics Miscellany and Puzzles:
http://www.cut-the-knot.org/m/Algebra/InequalityWithThreeConstraints.shtml

Martin Celli. Departamento de Matemáticas, Universidad Autónoma Metropolitana-Iztapalapa. Ciudad de México. E-mail: cell@xanum.uam.mx