Course notes for MATH 524: Non-Linear Optimization

Francisco Blanco-Silva

Department of Mathematics, University of South Carolina
E-mail address: blanco@math.sc.edu
URL: people.math.sc.edu/blanco

Contents

List of Figures

CHAPTER 1

Optimization Review from Vector Calculus

The starting point of these notes is the concept of *optimization* as developed in MATH 241 (see e.g. [**1**, Chapter 14])

DEFINITION. Let $D \subseteq \mathbb{R}^2$ be a region on the plane containing the point (x_0, y_0). We say that the real-valued function $f\colon D \to \mathbb{R}$ has a *local minimum* at (x_0, y_0) if $f(x_0, y_0) \leq f(x, y)$ for all domain points (x, y) in an open disk centered at (x_0, y_0). In that case, we also say that $f(x_0, y_0)$ is a *local minimum value* of f in D.

Emphasis was made to find conditions on the function f to guarantee existence and identification of minima:

THEOREM 1.1. *Let $D \subseteq \mathbb{R}^2$ and let $f\colon D \to \mathbb{R}$ be a function for which first partial derivatives $\frac{\partial f}{\partial x}$ and $\frac{\partial f}{\partial y}$ exist in D. If $(x_0, y_0) \in D$ is a local minimum of f, then $\nabla f(x_0, y_0) = 0$.*

The local minima of these functions are among the zeros of the equation $\nabla f(x, y) = 0$, the so-called *critical points* of f. More formally:

DEFINITION. An interior point of the domain of a function $f(x, y)$ where both directional derivatives are zero, or where at least one of the directional derivatives do not exist, is a *critical point* of f.

In order to select the minima, we employed the *Second Derivative Test for Local Extreme Values*

THEOREM 1.2. *Suppose that $f\colon \mathbb{R}^2 \to \mathbb{R}$ and its first and second partial derivatives are continuous throughout a disk centered at the point (x_0, y_0), and that $\nabla f(x_0, y_0) = 0$. Then $f(x_0, y_0)$ is a local minimum value if the two following conditions are satisfied:*

$$\frac{\partial^2 f}{\partial x^2}(x_0, y_0) > 0 \tag{1}$$

$$\det \underbrace{\begin{bmatrix} \frac{\partial^2 f}{\partial x^2}(x_0, y_0) & \frac{\partial^2 f}{\partial x \partial y}(x_0, y_0) \\ \frac{\partial^2 f}{\partial y \partial x}(x_0, y_0) & \frac{\partial^2 f}{\partial y^2}(x_0, y_0) \end{bmatrix}}_{\operatorname{Hess} f(x_0, y_0)} > 0 \tag{2}$$

REMARK 1.1. The restriction to univariate functions is even simpler: Suppose f'' is continuous on an open interval that contains x_0. If $f'(x_0) = 0$ and $f''(x_0) > 0$, then f has a local minimum at x_0.

EXAMPLE 1.1 (Rosenbrock Functions). Given strictly positive parameters $a, b > 0$, consider the (a, b)–Rosenbrock function

$$\mathcal{R}_{a,b}(x, y) = (a - x)^2 + b(y - x^2)^2.$$

It is easy to see that Rosenbrock functions are polynomials (prove it!). The domain is therefore the whole plane. Figure 1 illustrates a contour plot with several level lines of $\mathcal{R}_{1,1}$ on the domain $D = [-2, 2] \times [-1, 3]$, as well as its graph.

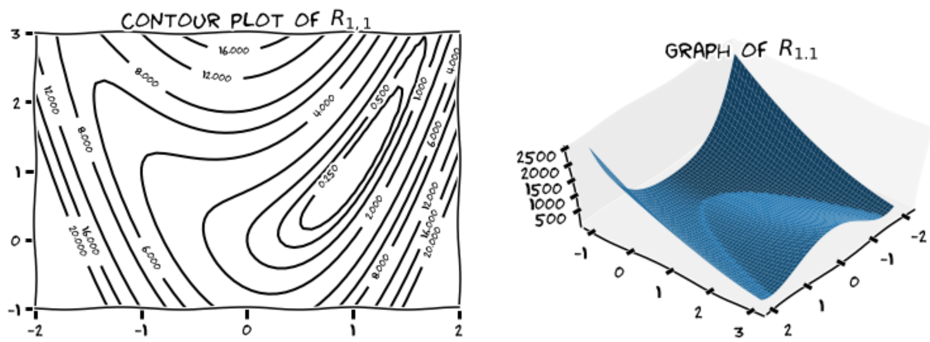

FIGURE 1. Details of the graph of $\mathcal{R}_{1,1}$

It is also easy to verify that the image is the interval $[0, \infty)$. Indeed, note first that $\mathcal{R}_{a,b}(x, y) \geq 0$ for all $(x, y) \in \mathbb{R}^2$. Zero is attained: $\mathcal{R}_{a,b}(a, a^2) = 0$. Note also that $\mathcal{R}_{a,b}(0, y) = a^2 + by^2$ is a polynomial of degree 2, therefore unbounded.

Let's locate all local minima:

- The gradient and Hessian are given respectively by

$$\nabla \mathcal{R}_{a,b}(x, y) = \big[2(x - a) + 4bx(x^2 - y), b(y - x^2)\big]$$

$$\operatorname{Hess}\mathcal{R}_{a,b}(x, y) = \begin{bmatrix} 12bx^2 - 4by + 2 & -4bx \\ -4bx & 2b \end{bmatrix}$$

- The search for critical points $\nabla \mathcal{R}_{a,b} = \mathbf{0}$ gives only the point (a, a^2). (Why?)
- $\frac{\partial^2 \mathcal{R}_{a,b}}{\partial x^2}(a, a^2) = 8ba^2 + 2 > 0$.
- The Hessian at that point has positive determinant:

$$\det \operatorname{Hess}\mathcal{R}_{a,b}(a, a^2) = \det \begin{bmatrix} 8ba^2 + 2 & -4ab \\ -4ab & 2b \end{bmatrix} = 4b > 0$$

There is only one local minimum at (a, a^2).

The second step was the notion of *global (or absolute) minima*: points (x_0, y_0) that satisfy $f(x_0, y_0) \leq f(x, y)$ for any point (x, y) in the domain of f. We always started with the easier setting, in which we placed restrictions on the domain of our functions:

THEOREM 1.3. *A continuous real-valued function always attains its minimum value on a* compact *set K. To search for global minima, we perform the following steps:*

- **Interior Candidates:** *List the critical points of f located in the interior of K.*
- **Boundary Candidates:** *List the points in the boundary of K where f may have minimum values.*
- **Evaluation/Selection:** *Evaluate f at all candidates and select the one(s) with the smallest value.*

EXAMPLE 1.2. A flat circular plate has the shape of the region $x^2+y^2 \leq 1$. The plate, including the boundary, is heated so that the temperature at the point (x, y) is given $T(x, y) = x^2 + 2y^2 - x$. Find the temperature at the coldest point of the plate.

We start by searching for critical points. The equation $\nabla T(x, y) = 0$ gives $x = \frac{1}{2}$, $y = 0$. The point $\big(\frac{1}{2}, 0\big)$ is clearly inside of the plate. This is our first candidate.

The border of the plate can be parameterized by $\varphi(t) = (\cos t, \sin t)$ for $t \in [0, 2\pi)$. The search for minima in the boundary of the plate can then be coded as an optimization problem for the function $f(t) = T \circ \varphi(t) = \cos^2 t + 2\sin^2 t - \cos t$ on the interval $[0, 2]$. Note that $f'(t) = 0$ at $t \in \{0, \frac{2}{3}\pi\}$ in $[0, 2\pi]$. We thus have two more candidates:

$$\varphi(0) = (1, 0) \qquad \varphi\big(\tfrac{2}{3}\pi\big) = \big(-\tfrac{1}{2}, \tfrac{1}{2}\sqrt{3}\big)$$

Evaluation of the function at all candidates gives us the answer: $T\big(\frac{1}{2}, 0\big) = -0.25$.

On a second setting, we do not place any restriction on the domain of the function. In this case, global minima will only be guaranteed with restrictions on the function:

DEFINITION (Coercive functions). A continuous real-valued function f is said to be *coercive* if for all $M > 0$ there exists $R = R(M) > 0$ so that $f(\boldsymbol{x}) \geq M$ if $x^2 + y^2 \geq R^2$. In other words:

$$\lim_{\|\boldsymbol{x}\| \to \infty} f(\boldsymbol{x}) = +\infty$$

EXAMPLE 1.3. Any polynomial $p_n(x) = a_n x^n + a_{n-1}x^{n-1} + \cdots + a_0$ with even degree $n \geq 2$ and positive leading coefficient is trivially coercive. Indeed; notice that we may write

$$a_n x^n + a_{n-1}x^{n-1} + \cdots + a_0 = a_n x^n \big(1 + \tfrac{a_{n-1}}{a_n x} + \cdots + \tfrac{a_0}{a_n x^n}\big)$$

The behavior of each of the factors as the absolute value of x goes to infinity leads to the statement.

$$\lim_{|x| \to \infty} a_n x^n = +\infty,$$

$$\lim_{|x|\to\infty} \left(1 + \tfrac{a_{n-1}}{a_n x} + \dots + \tfrac{a_0}{a_n x^n}\right) = 1.$$

In two dimensions, we must be careful assessing coerciveness of polynomials. Notice for example $p_2(x, y) = x^2 - 2xy + y^2$. Note how $p_2(x, x) = 0$ for any $x \in \mathbb{R}$, which proves p_2 is not coercive.

To see that the polynomial $p_4(x, y) = x^4 + y^4 - 3xy$ is coercive, we start by factoring the leading terms:

$$x^4 + y^4 - 3xy = (x^4 + y^4)\left(1 - \frac{3xy}{x^4 + y^4}\right)$$

Assume $r > 1$ is large, and that $x^2 + y^2 = r^2$. We have then

$$x^4 + y^4 \geq \frac{r^4}{2} \qquad \text{(Why?)}$$

$$|xy| \leq \frac{r^2}{2} \qquad \text{(Why?)}$$

therefore,

$$\frac{3xy}{x^4 + y^4} \leq \frac{3}{r^2}$$

$$1 - \frac{3xy}{x^4 + y^4} \geq 1 - \frac{3}{r^2}$$

$$(x^4 + y^4)\left(1 - \frac{3xy}{x^4 + y^4}\right) \geq \frac{r^2(r^2 - 3)}{2}$$

We can then conclude that given $M > 0$, if $x^2 + y^2 \geq \frac{1}{2}\big(3 + \sqrt{9 + 8M}\big)$, then $f(x, y) \geq M$.

Exercises

Problem 1.1. Develop similar statements as in Definition 1, Theorems 1.1, 1.2 and 1.3, but for *local* and *global maxima*.

Problem 1.2 (Domains). Find and sketch the domain of the following functions.

(a) $f(x, y) = \sqrt{y - x - 2}$
(b) $f(x, y) = \log\big(x^2 + y^2 - 4\big)$
(c) $f(x, y) = \frac{(x-1)(y+2)}{(y-x)(y-x^3)}$
(d) $f(x, y) = \log(xy + x - y - 1)$

Problem 1.3 (Contour plots). Find and sketch the level lines $f(x, y) = c$ on the same set of coordinate axes for the given values of c.

(a) $f(x, y) = x + y - 1$, $c \in \{-3, -2, -1, 0, 1, 2, 3\}$.
(b) $f(x, y) = x^2 + y^2$, $c \in \{0, 1, 4, 9, 16, 25\}$.
(c) $f(x, y) = xy$, $c \in \{-9, -4, -1, 0, 1, 4, 9\}$

Problem 1.4. Use a Computer Algebra System of your choice to produce contour plots of the given functions on the given domains.

(a) $f(x,y) = (\cos x)(\cos y)e^{-\sqrt{x^2+y^2}/4}$ on $[-2\pi, 2\pi] \times [-2\pi, 2\pi]$.
(b) $g(x,y) = \dfrac{xy(x^2-y^2)}{x^2+y^2}$ on $[-1,1] \times [-1,1]$
(c) $h(x,y) = y^2 - y^4 - x^2$ on $[-1,1] \times [-1,1]$
(d) $k(x,y) = e^{-y} \cos x$ on $[-2\pi, 2\pi] \times [-2, 0]$

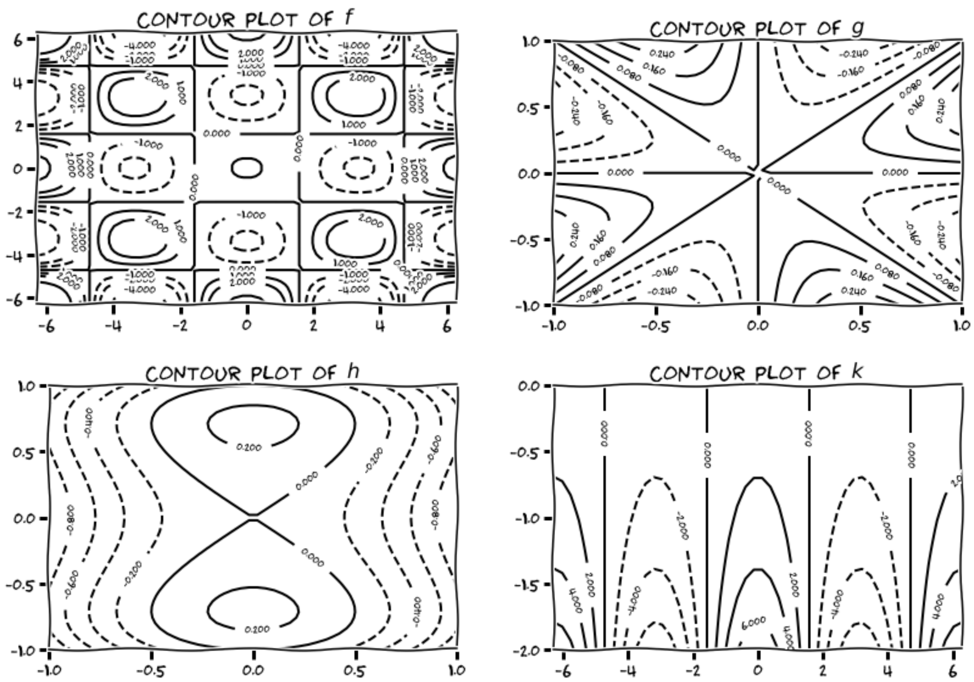

FIGURE 2. Contour plots for problem 1.4

PROBLEM 1.5. Find the points of the hyperbolic cylinder $x^2 = z^2 - 1 = 0$ in $\mathbb{R}^3$ that are closest to the origin.

PROBLEM 1.6. Prove that a coercive function always has a global minimum.

PROBLEM 1.7. Identify which of the following real-valued functions are coercive. Explain the reason.
(a) $f(x,y) = \sqrt{x^2+y^2}$.
(b) $f(x,y) = x^2 + 9y^2 - 6xy$.
(c) Rosenbrock functions $\mathcal{R}_{a,b}$.

PROBLEM 1.8. Find an example of a continuous, real-valued, non-coercive function $f \colon \mathbb{R}^2 \to \mathbb{R}$ that satisfies, for all $t \in \mathbb{R}$,

$$\lim_{x\to\infty} f(x,tx) = \lim_{y\to\infty} f(ty,y) = \infty$$

Bibliography

[1] Ross L Finney, Maurice D Weir, and George Brinton Thomas. *Thomas' calculus: early transcendentals*. Addison-Wesley, 2001.

[2] Anthony L Peressini, Francis E Sullivan, and J Jerry Uhl. *The mathematics of nonlinear programming*. Springer-Verlag New York, 1988.