THE CHINESE UNIVERSITY OF HONG KONG
Department of Mathematics
2017-2018 semester 2 MATH2010
week 11 tutorial

This week we continue with optimization theory. New mathematical tools in the our collection are first reviewed. Then we revisit old examples with new perspectives.

1 Theory

Definition 1. A symmetric matrix is $\left\{\begin{matrix}\text{positive definite}\\\text{negative definite}\\\text{indefinite}\end{matrix}\right\}$ iff its eignvalues $\left\{\begin{matrix}\text{are all positive}\\\text{are all negative}\\\text{have strictly opposite signs}\end{matrix}\right\}$.

Proposition 2. Consider a 2×2 *symmetric* matrix A.
A is $\left\{\begin{matrix}\text{positive or negative definite}\\\text{indefinite}\end{matrix}\right\}$ iff $\det A \left\{\begin{matrix}> 0\\< 0\end{matrix}\right\}$.
If $\det A > 0$, then A is $\left\{\begin{matrix}\text{positive definite}\\\text{negative definite}\end{matrix}\right\}$,
iff at least one diagonal entry is $\left\{\begin{matrix}\text{positive}\\\text{negative}\end{matrix}\right\}$.

Theorem 3. If the Hessian at an interior critical point is $\left\{\begin{matrix}\text{positive definite}\\\text{negative definite}\\\text{indefinite}\end{matrix}\right\}$, then it is a $\left\{\begin{matrix}\text{local minimum}\\\text{local maximum}\\\text{saddle point}\end{matrix}\right\}$

Theorem 4 (Lagrange multiplier – single-constrain)**.** Let f and g be C^1 real functions on $\Omega \subseteq \mathbb{R}^d$. Let x^* be an (interior) local extremum of f in $\{x : g(x) = 0\}$. If $\{\nabla g(x^*)\}$ is linearly independent, or equivalently, $\nabla g(x^*) \neq 0$, then $\nabla f(x^*) = \lambda \nabla g(x^*)$ for some λ.

The assumption that $\nabla g(x^*) \neq 0$ is crucial. Consider the following example.

Example 5. Let $g(x, y) = x^2$ and $f(x, y) = e^x + y^2$. Consider the problem $\min_{g(x,y)=0} f(x, y)$. This problem has a unique local minimum that has no Lagrange multiplier.

Proof. exercise. □

2 Examples

Example 6. The following table shows the second derivatives of some critical points of a C^2 function on $\mathbb{R}^2$. Classify the critical points.

critical points	f_{xx}	f_{xy}	f_{yy}
A	1	2	3
B	2	1	3
C	-1	2	3
D	-2	1	-3
E	1	3	9

Solution. Refer to the following table.

critical points	f_{xx}	f_{xy}	f_{yy}	H	$\det H$	extra checking	conclusion
A	1	2	3	$\begin{bmatrix}1 & 2\\ 2 & 3\end{bmatrix}$	-1	N/A	saddle point
B	2	1	3	$\begin{bmatrix}2 & 1\\ 1 & 3\end{bmatrix}$	5	$f_{xx} > 0$	local minimum
C	-1	2	3	$\begin{bmatrix}-1 & 2\\ 2 & 3\end{bmatrix}$	-7	N/A	saddle point
D	-2	1	-3	$\begin{bmatrix}-2 & 1\\ 1 & -3\end{bmatrix}$	5	$f_{xx} < 0$	local maximum
E	1	3	9	$\begin{bmatrix}1 & 3\\ 3 & 9\end{bmatrix}$	0	N/A	inconclusive

Remark. The eigenvalues of the Hessian are as follows.

critical point	eigenvalues of Hessian
A	$2+2\sqrt{2}$ (positive) and $2-2\sqrt{2}$ (negative)
B	$\frac{5\pm\sqrt{5}}{2}$ (positive)
C	$-1+\sqrt{29}$ (positive) and $-1-\sqrt{29}$ (negative)
D	$\frac{-5\pm\sqrt{5}}{2}$ (negative)
E	10 (positive) and 0

Example 7 (AM-HM inequality – reprise)**.** Suppose $x_i > 0$. Then $\frac{n}{\sum \frac{1}{x_i}} \leq \frac{1}{n}\sum x_i$.

Proof. Let $A(x_i) = \frac{1}{n}\sum x_i$ and $H(x_i) = \frac{n}{\sum \frac{1}{x_i}}$. (In this proof, H always means the harmonic mean. The Hessian is not used in this proof.)

It suffices to show that for $M > 0$,

$$M = \min_{\substack{H(x_i)=M \\ x_i>0}} A(x_i)$$

For interior local minima, $\nabla A = \frac{1}{n}[1]$ and $\nabla H = \frac{H^2}{n}[\frac{1}{x_i^2}]$. By the method of multipliers, since $\nabla H \neq 0$ (neither $\frac{H}{n}$ nor $\frac{1}{x_i^2}$ can be zero), $\frac{1}{n} = \lambda \frac{M^2}{n}\frac{1}{x_i^2}$, and hence $x_i = \sqrt{\lambda}M$ (the equation forces $\lambda > 0$ because everything else is positive). Plugging in $H(x_i) = M$ gives $\lambda = 1$, and hence $x_i = M$ and $A(x_i) = M$. Denote this point by $x^* = (M, ..., M)$.

Note that the domain is open, so there is no boundary local minimum. However, since it is neither closed nor bounded, the existence remains of a global minimum has to be established by restricting to a closed and bounded set.

Consider the region $R_B = \{(x_i) : \frac{1}{B} \leq x_i \leq B, H(x_i) = M\}$, which is closed and bounded and contains x^* as an interior point for every $B > M$. Since A is continuous, it has a minimum in R_B. Outside of R_B, either $x_j < \frac{1}{B}$ or $x_j > B$ for some j. The former is impossible whenever $B > \frac{M}{n}$, because $\frac{1}{x_j} < \sum \frac{1}{x_i} = \frac{n}{H(x_i)} = \frac{n}{M}$. For the latter, $A(x_i) > \frac{B}{n}$, which is larger than M whenever $B > Mn \geq \frac{M}{n}$. Then the minimum on R_B for $B = Mn + 1$ is the minimum for the whole domain, and hence A indeed has a minimum on the domain.

This global minimum has to be an interior local minimum, which by the method of multipliers is M. The result then follows.

□

Remark. This problem can also be cast as an unconstrained optimization problem, for instance $0 = \min A(x_i) - H(x_i)$. However, such problems have more complicated near-infinity and/or near-boundary behavior. This can be seen from the fact they have a continuum of minima, namely all x_i are equal.

Remark. One student conjectured that a unique critical point that is a local minimum of a C^1 function is a global minimum. This is NOT true. A counterexample can be found here: `https://math.stackexchange.com/questions/121326/unique-critical-point-does-not-imply-global-maximum-global-minimum`.

Example 8 (Maximum entropy distribution – reprise)**.** Let $f(p_i) = \sum p_i \log p_i$. Solve

$$\min_{\substack{\sum p_i = 1 \\ p_i \geq 0}} f(p_i)$$

Solution. This problem was shown to have a solution in week 9. This time, more sophisticated tools are used to find the solution.

For interior local minima, the method of multipliers gives $1 + \log p_i = \lambda$. In particular, all p_i's are equal, and hence $p_i = 1/n$. Let $p^* = (1/n)$. $f(p^*) = \log n$

To show that the minimum is not attained on the boundary, it suffices to show the function is strictly convex[1] upon restriction to the segment between each boundary point and the local minimum p^*.

Let p_0 be a boundary point and $\varphi(t) = f(p^* + t(p_0 - p^*))$. Since $\varphi''(t) = (p_0 - p^*)^t H(p_0 - p^*)$, so to show $\varphi''(t) > 0$ for $0 < t < 1$, it suffices to show H is positive definite. Indeed $H = \text{diag}(1/p_i)$ is positive definite. The conclusion then follows.

3 Epilogue

This was not part of the tutorial. It is only for your own reference.

[1] "Convex" always means "concave up".

The course is coming to an end, but the materials of this course is only the tip of the iceberg, both theoretically and from the viewpoint of application.

Theoretically, norm, on which virtually all limiting arguments in the course are based, is only a special instance of the more general concept of **metric**, for which Rudin's *Principles of Mathematical Analysis* is a good reference. Even more generally, our discussion of open sets, closed sets and closed and bounded sets is a precursion to the theory of **topology**, for which Munkre's *Topology* is a good starting point. To deepen our brief treatment of geometric objects in $\mathbb{R}^n$, the concept of **manifold** is indispensable, and Do Carmo's *Differential Geometry of Curves and Surfaces* and Lee's *Introduction to Smooth Manifolds* are standard texts on this subject.

Of course, an obvious sequel to this course is the study of **multivariable integral calculus**, which may be accompanied by Spivak's *Calculus on Manifolds*. Multivariable integral calculus is essential to the study of **partial differential equations** (e.g. the Laplace's equation $\Delta u = \sum D_{ii}u = 0$), which was briefly discussed in the tutorial, just as integration is essential for solving ordinary differential equation. A good introduction to the theory of partial differential equations would be Evan's *Partial Differential Equations*.

From the point of view of application, the study of partial differential equations in the last paragraph is a field with wide-ranging real-life application, as they can be used to model the financial market, weather and species population, to name a few applications. **Optimization** is also an important mathematical tool in many disciplines for inducing optimal performance of a system, or for modelling natural phenomena. A good reference to the mathematical theory optimization would be Rockafellar's *Variational Analysis*.

Finally, good luck with the examination. Feel free to make an appointment with me for a mathematical discussion. I wish you every success in your pan-mathematical career.