Comments on some chapters in "Mathematics for Economics" (third edition, 2011) by M. Hoy et al.

Erik Balder, Mathematical Institute, University of Utrecht, Netherlands

Fall 2012

1 Introduction

In 2010 a sudden departure at the Economics group of Utrecht University caused one of their courses to be without its regular teacher. This course is about Mathematical Economics and uses the textbook "Mathematics for Economics" by M. Hoy, J. Livernois, C. McKenna, R. Rees and T. Stengos (abbreviated as MfE below). So in 2010 the Economics group turned to the Mathematical Institute in Utrecht for help, and this caused me to be the teacher ad interim of their course.

The present note collects some of my findings about the MfE. They are limited to the chapters that I treated in 2010 and to those annotations of mine about them that I could still retrace. In the meantime, a third edition of MfE has appeared (2011, MIT Press). I was both surprised and disappointed to see that most of my points of criticism apply equally strongly to this later edition, but this does explain why I wrote the present note almost two years after having taught from MfE. I hope that MfE's authors will understand that the criticism in this note is meant to be constructive, in spite of its severe nature. So it is my sincere wish that future editions will live up to what is said in the preface about MfE's mission: *"All this must be done without sacrificing anything in terms of the rigor and correctness of the mathematics itself"* (p. xiv).

My homepage `http://www.staff.science.uu.nl/~balde101` contains some other relevant educational material; in particular, I refer to my critical evaluation of optimization methods used in the advanced microeconomics literature in "Exact and useful optimization methods for microeconomics", Springer Lecture Notes on Economics and Mathematical Systems 655 (2012), pp. 21-38, a paper of mine that can be downloaded there.

2 Points of serious concern

a. A considerable piece of educational material is being missed in MfE, causing at least one of its own exercises to be out of order. It is as follows. In contrast to the treatment of the Lagrange method in Chapter 13, the treatment of the Kuhn-Tucker theorem (Ch. 15) is only presented for concave programming problems (with inequality constraints only). Thus, for instance, the non-concave "Kuhn-Tucker-in-statu-nascendi" material in dimension one (section 12.3) remains without its logical multivariate follow-up in MfE. A powerful tool is thus completely missed: e.g., think of maximizing a non-concave utility function, such as the important Cobb-Douglas function $x^\alpha y^\beta$ for $\alpha+\beta>1$, over the consumer's budget set. For this reason Exercise 1 of section 15.2 (p. 578), where the Cobb-Douglas function to be maximized under inequality constraints is precisely of this non-concave kind, is out of order (all the more so since MfE does not seem to pay attention to equivalence of utility functions – thus, another most useful tool for the student is missing as well).

b. At least one theorem in section 13.3 would seem to be false and another theorem would seem incomprehensible at best:

Counterexample against Theorem 13.7. Consider in dimension one the function $f:\mathbb{R}\to\mathbb{R}$, given by $f(x):=(x+1)^3$ on $(-\infty,-1)$, $f(x):=0$ on $[-1,1]$ and $f(x):=(x-1)^3$ on $[1,+\infty)$. This function is quasiconcave. Now consider the null function $g^1(x)=0$, which is quasiconvex. According to Theorem 13.7 the point $x=0$, which is a locally optimal solution for maximization of $f(x)$ over all x with $g^1(x)=0$ (i.e., over all $x\in\mathbb{R}$), would be a *global* optimal solution for maximizing f over $\mathbb{R}$, but this is not true (incidentally, note that the same counterexample continues to hold if the constraint $g^1(x)=0$ is relaxed into $g^1(x)\leq 0$, a more natural type of constraint for the quasiconvex g^1).

Counterexample against Theorem 13.8(i). I could not track down the definition of "increasing" for a function of several variables in MfE (it seems to be missing in its index, although the concept of a "monotonically increasing sequence" was included). Yet by acceptable standards the functions $g^1(x_1,x_2):=\max(x_1,x_2)-2$ and $g^2(x_1,x_2):=x_1+x_2-3$ can be called "increasing" in that they satisfy $x_1'>x_1,\ x_2'>x_2\Rightarrow g^j(x_1',x_2')>g^j(x_1,x_2)$. Clearly, they are also quasiconvex. Now the set $S:=\{(x_1,x_2):g^1(x_1,x_2)=g^2(x_1,x_2)=0\}$ consists of the two isolated points $(2,1)$ and $(1,2)$. Consider any "increasing" and strictly quasiconcave function f with $f(2,1)>f(1,2)$. Then f has a local maximum at the isolated point $(1,2)$, but clearly this is not a global maximum.

By obvious modification, the above instance with two isolated points forming the set S can also be turned into a counterexample against part *ii* of Theorem 3.8.

Both counterexamples point at a problem that any critical student of the book should experience: claims, theorems, etc. (and apparently also adjectives like "increasing") are pronounced in an ex cathedra manner, without any hints or references to bolster up their proofs.

c. A major error on p. 519: the scope of the second order conditions for optimality, which only hold *at the particular point* being inspected on p. 519 for optimality (i.e., locally), is incorrectly turned into the global scope that is needed for quasiconcavity.

d. On p. 523 the first five lines claim something that is said to be *"straightforward"*, but is, in fact, quite wrong and most misleading for the student: the claim would only hold if the functions f and g^i were concave and if inequality constraints were used (giving the multipliers the appropriate sign), but that is not the case at all on p. 523. Here is an obvious counterexample:

Counterexample against the statement contained in the first five lines of p. 523. For $n = 1$ let $f(x) := x^3$ and let $g^1(x) := x$. Then $S := \{x : g^1(x) = 0\} = \{0\}$, so the unique optimal solution is $x^* = 0$ (note that $g_1^1 = 1$, so the full rank condition is fulfilled). Solving $3 * 0^2 + \lambda_1 * 1 = 0$ gives $\lambda_1 = 0$. However, it is not true that $0 = f(x^*) + 0 * g^1(x^*) \geq x^3 + 0 * g_1(x) = x^3$ holds for all x.

e. In several places in MfE the text offers possibilities for confusion about local versus global optimality. Clearer language seems preferable, all the more since obfuscation has sometimes led to errors in logical thinking as well. For instance, "solves" in Theorem 13.4 means "locally solves". Another instance is Example 13.3, which starts out by stating that it will produce what is remarkably called *"a true maximum"* (a remnant from the second edition, where this term was used a lot in similar instances). But the truth of the matter is that the last line of Example 13.3 only allows the student to conclude – at that stage – that "we have a maximum" only means "we have a *local* maximum". Thereupon, to achieve what is truly a maximum (i.e., a global maximum), the example should have been continued. Namely, by different means and reasoning of a global nature (either based on the Weierstrass Theorem 13.5 or on the concavity of the utility function that is maximized), one can show that this is actually also a global maximum. A similar comment applies to the ancillary Example 13.1.[1]

f. Examples, such as Examples 13.1, 13.2 and 13.3, and exercises, such as Exercises 1-2 in section 13.1 or Review Exercise 1(f) in Ch. 12 (but there are many more instances, including in other chapters, such as Ch. 25), raise fundamental questions about the *domain of definition* of the functions on which the various rules have been applied. For instance, the unsuspecting student of the more theoretical material in Chapter 13 could easily fall into thinking that the variables x_1 and x_2 can take any value, i.e., that optimization is performed with functions f and g^i that are defined on all of $\mathbb{R}^2$, because the contrary seems to be mentioned nowhere in the text. Yet the examples and exercises sometimes refer implicitly to a different theoretical framework, where variables must be strictly positive or nonzero, etc. As long as such domains are open, the required adjustments can be made with relative ease. However, often those domains are actually closed in economics: for instance, Cobb-Douglas-, CES-, Leontiev- and linear utility functions are all defined for what are *nonnegative* variables in principle. One sees this kind of deficiency all over the economic place, but this is a fine example of a situation where economists cannot simply rely on or copy what mathematicians usually do: see my downloadable paper which I mentioned in the introduction.

[1] Incidentally, the Hessian H calculated in Example 13.3 has some errors.

g. It would have been very useful, if not essential (e.g., see the previous observation about Example 13.3), to add in Theorem 12.4 that if f is merely concave, then its stationarity condition still implies that $\mathbf{x}^*$ is a global maximum.

h. P. 850: the last line is false. Maximizing a strictly concave function does not *imply* that the optimal solution should be interior (counterexample: consider $-x^2$ being maximized over $\mathbb{R}_+$ ...). Moreover, because domains are never mentioned explicitly (see my point f above), a student might well ask: *interior to what*? At any rate, it should be clear that the two assumptions (i.e., interiority and strict concavity) are really independent; together they imply what is desired on p. 850, i.e., simply setting the derivative in y equal to zero is enough.

i. P. 900 could either point at a serious misunderstanding or at a typo. However, because $y < h(x, y, t)0$ (which should be $y \leq h(x, y, t)$ there) is explained by *"As usual, these conditions must hold with complementary slackness. If $\theta = 0$ the constraint is not binding and we have $y < h(x, y, t)$."*, it seem suspiciously close to a real misunderstanding. Fortunately, to a certain degree the attentive student can already deduce from p. 493 (which deals with dimension 1) what complementary slackness means when a multiplier is zero. Incidentally, I could not find any trace of the term "complementary slackness" in the index of MfE.

j. P. 880: Example 25.5 shows another misunderstanding about complementary slackness. Because it works with an inequality constraint, an additional sign condition for $\lambda(T)$ ought to be included in the necessary conditions for optimality in case 2: it should be nonnegative (this is not unlike a similar condition for the multipliers in the Kuhn-Tucker theorem). Incidentally, there is a typo in the last line of this example, where Example 25.5 should be replaced by Example 25.3.

3 Smaller points

a. For a book of almost 1000 pages, cutting out redundancies should be a major goal. One way to do this is to use the so-called "sign trick" by which solutions/results about positivity/maximization questions can be immediately turned into their counterparts for negativity/minimization questions and vice versa. For instance: a matrix A is positive definite [resp. positive semidefinite] if and only $-A$ is negative definite [resp. negative semidefinite].[2], a function f is convex [resp. strictly convex] if and only $-f$ is concave [resp. strictly concave], minimizing a function f over a set S gives the same optimal solutions as maximizing $-f$ over S, etc, etc. Once this is explained to the student, a lot of material becomes simply redundant. In this connection, the lines just before

[2]Observe that using this insight the memory-burden for the student becomes considerably lighter by trading in Theorem 10.19 for Theorem 10.18.

Theorem 12.2 on p. 477 or Theorem 12.8 on p. 493, etc. certainly do not place the student on the right track.

b. The reliance on differentials in many arguments in MfE is not a wise choice in my opinion. I am aware of the fact that it is used in other places in the mathematical economics literature as well, but I would like to point out that it mainly rests on purely symbolic formalism (coming from the theory of differential forms), for which better substitutes are available, certainly given MfE's readership.

c. By Definitions 9.4 and 10.13 a minor is a determinant of a matrix, but on p. 454 the student is told that a minor is a matrix.

d. A missed opportunity in the important Cobb-Douglas Example 11.31: apart from what is correctly mentioned for $\alpha + \beta < 1$, in the case $\alpha + \beta = 1$ the function is still concave and in the case $\alpha + \beta > 1$ it is still quasi-concave. Apart from this, wouldn't an economist-student be interested to learn which of these (quasi)concavity properties continue to hold for $x_1, x_2 \geq 0$ instead of $x_1, x_2 > 0$?

e. P. 486, middle: a point, such as $\mathbf{x}^*$ is not a value, so to say that *"$\mathbf{x}^*$ is an extreme value"* is confusing for the student. Compare this to the first lines of 12.1. Similar confusing terminology is used on p. 474: *"that a stationary value may be a* **saddle point***"*.

f. P. 516, towards bottom: *" in terms of small deviations around the optimal point"*. It might help the student, so as to better understand the borders of the Hessian, if this were stated as *" in terms of small deviations around the optimal point* **while maintaining feasibility***"*.

g. P. 554: there is a typo (minus sign instead of plus sign) in the formula in the middle of that page.

h. P. 418: typo: noninfinitessimal should be noninfinitesimal

i. On p. 609: why choose those curves to be linear in the figure?

j. A typo in Theorem 20.4: the provision should be $a_1 + a_2 \neq -1$.

k. Typos in Example 24.3: $\bar{y}_1$ is meant instead of y_1, etc.

l. A typo in Example 24.4: "solved in Example 24.3" should be "solved in Example 24.1".