# Exact and useful optimization methods for microeconomic theory

Erik J. Balder, Mathematical Institute, University of Utrecht,
P.O. Box 80.010, 3508 TA Utrecht, Netherlands

May, 2008

Mailing address: Mathematical Institute, University of Utrecht, Budapestlaan 6, P.O. Box 80.010, 3508 TA Utrecht, the Netherlands
email: E.J.Balder@uu.nl, fax: +31-30-2518394
JEL classification: D00

**Abstract.** This paper points out that the treatment of utility maximization in current textbooks on microeconomic theory is deficient in at least three respects: breadth of coverage, completeness-cum-coherence of solution methods and mathematical correctness. Improvements are suggested in the form of a Kuhn-Tucker type theorem that has been customized for microeconomics. The role of the domain of differentiability of the utility function is emphasized, not only to point out a persistent error in the standard literature on microeconomic theory, but also, more constructively, to call attention to the fact that this domain can be chosen sensibly in order to include the maximization of certain nondifferentiable utility functions, such as Leontiev utility functions. To ensure uniqueness of the optimal solution $(S)$-*quasiconcavity*, an apparently new adaptation of the notion of strict quasiconcavity, is introduced. It improves upon an earlier notion formulated by Aliprantis, Brown and Burkinshaw. To underscore the usefulness of the optimality conditions obtained here, five quite different instances of utility maximization are completely solved by a single coherent method.

## 1 Introduction

In the currently popular textbooks on microeconomic theory and mathematical economics [8, 10, 11, 13], all of which stress rigor and precision, the treatment of the fundamental subject of utility maximization would seem to show deficiencies in the following three respects: breadth of coverage, completeness-cum-coherence of solution methods and mathematical correctness. Similar deficiencies also show up in related textbooks with a less formal orientation [9, 14]. The net effect of this is that the reader will search in vain in the standard literature on microeconomic theory and mathematical economics for a method to derive the Marshallian demand function that meets the following criteria: it must be coherent and complete, based on generally accepted principles of correct mathematical reasoning (starting from, say, the Kuhn-Tucker and Weierstrass' theorems), and it must be operationally useful by being applicable to at least the following five instances of standard utility functions on the standard consumption set $\mathbb{R}^\ell_+$, where $\ell \in \mathbb{N}$ stands for the number of commodities: ($i$) Cobb-Douglas utility function, $\ell > 2$, ($ii$) CES utility function, $\ell > 2$, ($iii$) linear utility function with positive coefficients, $\ell \geq 2$, ($iv$) the utility function $u(x_1, x_2) = x_1^2(x_2 + 1)$ or any similar one leading to corner point solutions, and ($v$) Leontiev utility function, $\ell \geq 2$.

The already mentioned deficiencies in the standard literature are briefly discussed in Remark 1. I first observe that there are good reasons for including at least some instances with more than two commodities in the preceding list. Indeed, for $\ell = 2$ each of the above utility maximization problems can immediately be reduced to a scalar optimization problem over an interval; indeed, their utility functions are strictly increasing, causing any optimal solution to be budget-balanced by Theorem 1($a$). Such a scalar optimization problem is quite elementary: it can be completely solved by making sign diagrams of the derivative, possibly supported by the use of computer algebra packages. This explains why $\ell > 2$ was chosen in the classical instances ($i$)-($ii$) to test for operational usefulness of the method that I aim for. To apply to instances ($iii$)-($iv$) the solution method must be able to determine multiple solutions and corner point solutions (a standard subject in intermediate microeconomics courses [5]). Finally, instance ($v$) has a nondifferentiable utility function, which presents a quite unusual operational challenge.

In section 3 I shall develop an efficient method that meets all of the above criteria; it is based on optimality results for general utility maximization that are presented in section 2. This development started in my lecture notes [3]. Essentially, the optimality results take the form of a Kuhn-Tucker theorem, Theorem 1, that has been customized for microeconomics. To a high degree such customization depends on a special property of utility functions that is commonly found in microeconomics: they are strictly increasing. Beforehand, I observe that the classical Cobb-Douglas and CES instances ($i$)-($ii$) require that my model can deal with utility functions that are defined on the nonnegative orthant $\mathbb{R}^\ell_+$ only, with possible nondifferentiability on its boundary. Together with the applicability of my method to the nondifferentiable instance ($v$), this indicates that careful consideration of the differentiability domain of the utility function in the maximization problem plays an important part in this paper; cf. Remark 1. The customized Kuhn-Tucker theorem comes close to what is done on [8, p. 23 ff.], but, as my applications to instances ($i$), ($ii$) and ($iv$) will show in particular, it is part ($d$) of Theorem 1, which is absent in [8], that makes a considerable difference: it allows my method, in its handling of sufficient conditions for optimality and uniqueness, to go beyond interior optimal solutions. For example, this makes it possible to derive corner point solutions in a rigorous, coherent and efficient way in instance ($iv$), whose utility function is *not* strictly quasiconcave on $\mathbb{R}^2_+$. This part ($d$) uses an apparently new adaptation of the notion of strict quasiconcavity, called ($S$)-*quasiconcavity*, which improves upon a related earlier notion by Aliprantis, Brown and Burkinshaw [2] (see Remark 2).

In sum, this paper improves on the usual literature on utility maximization in microeconomics by presenting a Kuhn-Tucker theorem that exploits the usual strict monotonicity of utility functions in microeconomics via the new notion of ($S$)-quasiconcavity. The resulting solution method, which also brings out the role that the utility function's domain of differentiability can play, is both rigorous and versatile.

## 2 Customized optimality results for microeconomics

Let $u : \mathbb{R}^\ell_+ \to \mathbb{R}$ be a continuous function, the *utility function.* Let $\Omega$ be an open set that is contained in $\mathbb{R}^\ell_+$ (whence in $\mathbb{R}^\ell_{++}$); the function $u$ is supposed to be (Fréchet) differentiable on $\Omega$ [7]. In section 3 I shall choose for $\Omega$ the strictly positive orthant $\mathbb{R}^\ell_{++}$ so as to treat instances ($i$)-($iv$), but to deal with the Leontiev instance ($v$) I shall choose $\Omega$ differently. Throughout I suppose that $u$ is *strictly increasing* on $\mathbb{R}^\ell_+$; that is to say, for every $x$ and $x'$ in $\mathbb{R}^\ell_+$ the following must hold: if $x_i > x'_i$ for every $i = 1, 2, \ldots, n$, then $u(x) > u(x')$. For $p \in \mathbb{R}^\ell_{++}$ (*price vector*) and $y \in \mathbb{R}_+$ (*income*) the consumer's *utility maximization problem* is as follows:

$$\text{maximize } u(x) \text{ over all } x \in \mathbb{R}^\ell_+ \text{ such that } p \cdot x \leq y. \tag{1}$$

This problem is well-defined because $0 \in B$. Here $B := \{x \in \mathbb{R}^\ell_+ : p \cdot x \leq y\}$ stands for the feasible set of this problem, the *budget set.* I shall also use $B_0 := \{x \in \mathbb{R}^\ell_+ : p \cdot x = y\}$ to denote the so-called *budget plane.* A vector $x$ in $\mathbb{R}^\ell_+$ is said to be *budget-balanced* if $p \cdot x = y$, that is to say, if it belongs to the budget plane $B_0$. I shall say that the function $u : \mathbb{R}^\ell_+ \to \mathbb{R}$ has *property* ($S$) if for every pair $x, x' \in \mathbb{R}^\ell_+$ with $x \neq x'$ and $u(x) = u(x') > u(0)$ one has $u(\frac{1}{2}x + \frac{1}{2}x') > u(x) = u(x')$. The function $u : \mathbb{R}^\ell_+ \to \mathbb{R}$ is said to be ($S$)-*quasiconcave* if it is quasiconcave on $\mathbb{R}^\ell_+$ and has property ($S$). Note that ($S$)-quasiconcavity of $u$ is *always* considered on the full domain $\mathbb{R}^\ell_+$. It seems rather surprising that this natural modification of the classical notion of strict quasiconcavity, which excludes certain points that are sub-optimal (if $y > 0$) and only works with points $x$ and $x'$ at the same utility level, should be new; yet this appears to be the case. Clearly, a sufficient condition for $u$ to be $(S)$-quasiconcave is that it is quasiconcave on $\mathbb{R}^\ell_+$ and strictly quasiconcave when restricted to the set $\{x \in \mathbb{R}^\ell_+ : u(x) > u(0)\}$. A related and useful sufficient condition for $(S)$-quasiconcavity is given in Proposition 1 below, where I will also demonstrate that the utility function is $(S)$-quasiconcave in instances $(i)$, $(ii)$ and $(iv)$ .

**Theorem 1** $(a)$ *The consumer's utility maximization problem (1) has an optimal solution. Moreover, every optimal solution is budget-balanced.*

$(b)$ *Suppose that (1) has an optimal solution $x^*$ which is such that $x^* \in \Omega$. Then there exists $\lambda \geq 0$ such that*

$$\nabla u(x^*) = \lambda p. \tag{2}$$

$(c)$ *If $x^* \in \Omega$ is budget-balanced and such that (2) holds for some $\lambda > 0$, then $x^*$ is an optimal solution of (1), provided that $u$ is quasiconcave on $\mathbb{R}^\ell_+$.*

$(d)$ *If $u$ has property $(S)$, then (1) has a unique optimal solution. In particular, if $u$ is $(S)$-quasiconcave, then any budget-balanced $x^* \in \Omega$ for which (2) holds for some $\lambda > 0$, is the unique optimal solution of (1).*

This result has the familiar makeup of results in optimization theory: existence, followed by necessary conditions for optimality that are sharpened into sufficient conditions and even a uniqueness condition. Simple examples demonstrate that the above formulation is sharp; for instance, taking $\ell = 1$, $u(x) := (x-1)^3$, $y = 1$ and $p = 1$ shows that the possibility $\lambda = 0$ cannot be excluded in part $(b)$, etc.

**Lemma 1** *Suppose that $u$ is quasiconcave on $\mathbb{R}^\ell_+$. Then for every $x \in \Omega$ and $x' \in \mathbb{R}^\ell_+$*

$$u(x) \leq u(x') \text{ implies } \nabla u(x) \cdot (x' - x) \geq 0.$$

*Proof.* For $t \in [0,1]$ let $\phi(t) := u(tx + (1-t)x')$. Then $u(x) \leq u(x')$ implies $\phi(t) \geq u(x) = \phi(1)$. So $\phi(t)$ attains a minimum over $[0,1]$ for $t = 1$. Also, because $u$ is differentiable at $x \in \Omega$, the function $\phi$, which is the composition of $u$ and a linear mapping, is differentiable from the left at 1. These two facts imply that $\phi'(1) \leq 0$. The desired inequality then follows by the chain rule. QED.

*Proof of Theorem 1.* $(a)$ Existence of an optimal solution $x^*$ follows by the Weierstrass theorem, because $u$ is continuous and $B$ is a nonempty compact set. For $y = 0$ the identity $B = B_0 = \{0\}$ causes budget balancedness to hold trivially. For $y > 0$, $x^* \notin B_0$ would imply $p \cdot x^* < y$. So setting $\tilde{x}_i := x^*_i + t$ would result in a contradiction for $t > 0$ sufficiently small, because then $\tilde{x} \in B$ and $u(\tilde{x}) > u(x^*)$.

$(b)$ The hypothesis $x^* \in \Omega$ implies that $x^*$ is also an optimal solution of the auxiliary optimization problem

$$\text{maximize } u(x) \text{ over all } x \in \Omega \text{ with } p \cdot x \leq y, \tag{3}$$

which has only one inequality constraint. By the Kuhn-Tucker theorem in [13, Theorem 1.D.3], which allows open domains of definition for its functions instead of all of $\mathbb{R}^\ell$, it follows that there exists $\lambda \geq 0$ such that $\nabla u(x^*) = \lambda p$ (furthermore, it implies that $\lambda = 0$ if $x^* \notin B_0$, but this is irrelevant by part $(a)$). Here the constraint qualification in Theorem 1.D.3 is satisfied by applying Theorem 1.D.4 in [13].

$(c)$ Suppose that $x^* \in B_0 \cap \Omega$ would not be optimal. Then there would be $\tilde{x} \in B$ such that $u(\tilde{x}) > u(x^*)$. So $u((1-t)\tilde{x}) > u(x^*)$ would hold for $t > 0$ small enough, by continuity of $u$. By Lemma 1 the quasiconcavity hypothesis for $u$ implies

$$\nabla u(x^*) \cdot (x - x^*) \geq 0 \text{ for every } x \in \mathbb{R}^\ell_+ \text{ with } u(x) \geq u(x^*).$$

By $\nabla u(x^*) = \lambda p$ and $\lambda > 0$, it then follows that $p \cdot ((1-t)\tilde{x} - x^*) \geq 0$, whence $p \cdot \tilde{x} > (1-t)p \cdot \tilde{x} \geq y$ by the given budget-balancedness of $x^*$ (observe that $p \cdot \tilde{x} > 0$ by $\tilde{x} \neq 0$). This contradicts $\tilde{x} \in B$.

$(d)$ First, if $y = 0$ then $B = \{0\}$, so $x^* = 0$ is the unique optimal solution. Next, if $y > 0$, then $u(x^*) > u(0)$ must hold for any optimal solution $x^*$, because $(t, t, \ldots, t)$ belongs to $B$ for $t > 0$ small enough and because $u$ is strictly monotone. Now suppose that $x^*$ and $x^{**}$ were two different optimal solutions of (1). Then $u(x^*) = u(x^{**})$ = optimal value of (1) and $u(x^*) > u(0)$ by the previous argument. Define $\tilde{x} := \frac{1}{2}x^* + \frac{1}{2}x^{**}$; then $\tilde{x} \in B$ and property $(S)$ gives $u(\tilde{x}) > \frac{1}{2}u(x^*) + \frac{1}{2}u(x^{**}) = u(x^*)$. This contradicts the optimality of $x^*$. So the optimal solution is unique. The final part of the statement is an immediate consequence of combining part $(c)$ with uniqueness. QED

**Remark 1** An alternative application of the Kuhn-Tucker theorem is obtained if, instead of working with the above open set $\Omega \subset \mathbb{R}^\ell_+$, one uses a model where the utility function $u$ is defined and differentiable on an open set $\Omega'$ that *contains* $\mathbb{R}^\ell_+$. In that case optimality of $x^*$ for (1) can be expressed equivalently as optimality of $x^*$ for the following optimization problem:

$$\text{maximize } u(x) \text{ over all } x \in \Omega' \text{ with } p \cdot x \leq y \text{ and } -x_i \leq 0,\ i = 1, \ldots, \ell. \tag{4}$$

Precisely such an application was chosen in [9, 10, 11, 13], but in doing so one forms a model that no longer applies to all Cobb-Douglas or CES utility functions. In [9, p. 131 ff.], [10, p. 50 ff.], [11, Theorem 22.1, Example 22.1] and [14, section 2.2] (the latter reference discusses this for profit maximization) this has led to imprecise or incorrect formulations of their necessary first order optimality conditions and, in the case of [9, 10, 11], to an incorrect application to the Cobb-Douglas instance $(i)$.[1] Apart from this, the derivation of these optimality conditions is standard and can be found in [10, 11]: by applying the Kuhn-Tucker to (4), which has $\ell + 1$ inequality constraints, one now obtains as a first-order necessary condition for optimality

$$\frac{\partial u(x^*)}{\partial x_i} \leq \lambda p_i, \text{ with equality if } x_i^* > 0,\ i = 1, \ldots, \ell$$

instead of (2) and this holds for every optimal $x^*$ in $\Omega'$, a set that now includes the boundary of $\mathbb{R}^\ell_+$. For utility functions that fit into this model, *but only for those*, this formulation gives meaningful and somewhat sharper results. These remarks do not affect [8], whose optimality results presuppose that the optimal solution should belong to $\mathbb{R}^\ell_{++}$, as exemplified by the application to the Cobb-Douglas instance in exercise 1.20 of [8]. This is not enough to allow a treatment of the corner point solution instances $(iii)$-$(iv)$. It reflects a common shortcoming of the above references: except for [14, p. 57], none would seem to contain a completely solved instance of a corner point solution.

From part $(d)$ of Theorem 1 it is evident that $(S)$-quasiconcavity can help to solve the optimization problem (1), but it should be kept in mind that this is very much due to the hypothesis that the utility function $u$ is strictly increasing: property $(S)$ ignores bundles at the lowest utility level, which is $u(0)$. It is obvious that strict quasiconcavity of $u$ on $\mathbb{R}^\ell_+$, such as in the CES instance $(ii)$, implies its $(S)$-quasiconcavity on $\mathbb{R}^\ell_+$; however, the converse is not true, as will be demonstrated in Example 1$(i)$ for the Cobb-Douglas instance $(i)$. An obvious sufficient condition for $(S)$-quasiconcavity, which invariably works and is in terms of standard properties, is to require $u$ to be quasiconcave on $\mathbb{R}^\ell_+$ and strictly quasiconcave on $\{x \in \mathbb{R}^\ell_+ : u(x) > u(0)\}$.[2] For operational use I shall state a simple sufficient condition. To prepare for it, I observe that the properties of $u$ in Theorem 1 cause the range of $u$ to be an interval, namely $[u(0), u_\infty)$, where $u_\infty := \sup_{x \in \mathbb{R}^\ell_+} u(x)$ (this notation allows for the possibility that $u_\infty$ equals $+\infty$). To see this, define $\psi(t) := u(t, t, \ldots, t)$ for $t \geq 0$; this is a continuous, strictly increasing function on $\mathbb{R}_+$. Clearly, $u_\infty = \sup_{t \geq 0} \psi(t)$ and the supremum cannot be attained. Because for any $x \in \mathbb{R}^\ell_+$ the value $u(x)$ lies between $\psi(0)$ and $\psi(t)$ for some $t > 0$ sufficiently large (this follows by strict monotonicity and continuity of $u$), the intermediate value theorem can be invoked to finish the argument.

**Proposition 1** *If $\{x \in \mathbb{R}^\ell_+ : u(x) > u(0)\}$ is convex and if there exists a strictly increasing function $h : (u(0), u_\infty) \to \mathbb{R}$ such that the composition mapping $x \mapsto h(u(x))$ is strictly concave (or strictly quasiconcave) on $\{x \in \mathbb{R}^\ell_+ : u(x) > u(0)\}$, then $u$ is quasiconcave on $\mathbb{R}^\ell_+$ and strictly quasiconcave on $\{x \in \mathbb{R}^\ell_+ : u(x) > u(0)\}$. In particular, $u$ is $(S)$-quasiconcave.*

[1] Notwithstanding its general Kuhn-Tucker Theorem 1.D.3, reference [13] considers utility maximization only for utility functions defined on all of $\mathbb{R}^\ell$ (see its pp. 134-135); on pp. 223-224 this has resulted in an *ad hoc* solution of instance $(i)$.

[2] It should be kept in mind that the simple property $(S)$ originated in a course at the first year undergraduate level [3].

*Proof.* First, I prove that $u$ is quasiconcave. Let $\alpha \in \mathbb{R}$. If $\alpha \leq u(0)$, then $\{u \geq \alpha\} := \{x \in \mathbb{R}^\ell_+ : u(x) \geq \alpha\} = \mathbb{R}^\ell_+$. If $\alpha > u(0)$, then for any $x, x' \in \{u \geq \alpha\} \subset \{u > u(0)\}$, $x \neq x'$, and for any $t \in (0,1)$ the given strict concavity (or strict quasiconcavity) property implies $h(u(tx+(1-t)x')) > \min(h(u(x)), h(u(x))) \geq h(\alpha)$. So $u(tx+(1-t)x') > \alpha$ by strict monotonicity of $h$. For the set $\{u \geq \alpha\}$ I conclude that it equals $\mathbb{R}^\ell_+$ for every $\alpha \leq u(0)$ and that it is strictly convex for every $\alpha > u(0)$. Hence, $u$ is certainly quasiconcave on $\mathbb{R}^\ell_+$. The desired strict quasiconcavity of $u$ on $\{u > u(0)\}$ also follows from the previous conclusion. QED

**Example 1** (*i*) Utility functions $u : \mathbb{R}^\ell_+ \to \mathbb{R}$ of Cobb-Douglas type are given by $u(x) := \Pi_{i=1}^\ell x_i^{\alpha_i}$ with all $\alpha_i > 0$. They are $(S)$-quasiconcave, but $u(x_1, 0, 0, \ldots, 0) = 0$ shows that they are not strictly quasiconcave on $\mathbb{R}^\ell_+$. Proposition 1 will be used to show $(S)$-quasiconcavity. Observe first that the set $\{x \in \mathbb{R}^\ell_+ : u(x) > u(0)\} = \mathbb{R}^\ell_{++}$ is convex. On $(u(0), u_\infty) = (0, +\infty)$ I choose $h(t) := \log(t)$; this is a strictly increasing function, which gives $h(u(x)) = \sum_{i=1}^\ell \alpha_i \log(x_i)$ on $\mathbb{R}^\ell_{++}$. Each function $x \mapsto \log(x_i)$ is strictly concave on $\mathbb{R}^\ell_{++}$ (simply because $x_i \mapsto \log(x_i)$ is strictly concave on $(0, +\infty)$), so the function $x \mapsto h(u(x))$, being the sum of strictly concave functions, is strictly concave on $\mathbb{R}^\ell_{++}$. Therefore, $u$ is $(S)$-quasiconcave on $\mathbb{R}^\ell_+$ by Proposition 1.

(*ii*) Utility functions $u : \mathbb{R}^\ell_+$ of CES type are given by $u(x) := (\sum_{i=1}^\ell x_i^\rho)^{1/\rho}$ with $0 < \rho < 1$. Proposition 1 will be applied to show that such $u$ are $(S)$-quasiconcave and even strictly quasiconcave on $\mathbb{R}^\ell_+$. First, I observe that the set $\{x \in \mathbb{R}^\ell_+ : u(x) > u(0)\} = \mathbb{R}^\ell_+ \backslash \{0\}$ is convex (because this only leaves out the origin, my argument can immediately be adapted to imply that $u$ is even strictly quasiconcave on $\mathbb{R}^\ell_+$). On $(u(0), u_\infty) = (0, +\infty)$ I choose $h(t) := t^\rho$, a strictly increasing function; then $x \mapsto h(u(x)) = \sum_{i=1}^\ell x_i^\rho$ is strictly concave on $\mathbb{R}^\ell_+$ by repeating the reasoning in part (*i*) (namely, each function $x_i \mapsto x_i^\rho$ is strictly concave on $\mathbb{R}_+$, because $0 < \rho < 1$). So the conditions in Proposition 1 certainly hold; hence, $u$ is $(S)$-quasiconcave on $\mathbb{R}^\ell_+$.

(*iii*) Let $u : \mathbb{R}^2_+ \to \mathbb{R}$ be given by $u(x_1, x_2) = x_1^2(x_2 + 1)$, as in my instance (*iv*). This function is $(S)$-quasiconcave (however, $u(0, x_2) = 0$ for all $x_2 \geq 0$ shows that it is not strictly quasiconcave on $\mathbb{R}^2_+$). To see this, I observe first that the set $\{x \in \mathbb{R}^2_+ : u(x) > u(0)\} = \{(x_1, x_2) : x_1 > 0, x_2 \geq 0\}$ is convex. On $(u(0), u_\infty) = (0, +\infty)$ I choose $h(t) := \log(t)$, a strictly increasing function; then $x \mapsto h(u(x)) = 2\log(x_1) + \log(x_2 + 1)$ is evidently strictly concave on $\{(x_1, x_2) : x_1 > 0, x_2 \geq 0\}$ (repeat the reasoning in part (*i*)). Hence, $u$ is $(S)$-quasiconcave by Proposition 1.

The above example is intended to be of direct use for applications of Theorems 1 to instances (*i*), (*ii*) and (*iv*) of my list. It is well-known that more can be said. For instance, for $\sum_i \alpha_i \leq 1$ the utility function in part (*i*) of Example 1 is concave on $\mathbb{R}^\ell_+$ and in part (*ii*) the same is even true for any $\rho \in (0,1)$. In both cases, after having established quasiconcavity, one can apply results in [6] or [12] to derive such *ex post* concavity properties. See Remark 4 in [6, p. 123] or Theorem 2.5.3 in [12] and the examples which follow that theorem.

**Remark 2** In [2] Aliprantis, Brown and Burkinshaw consider utility maximization for a pure exchange consumer, but their analysis extends effortlessly to that for an ordinary consumer. For $\ell = 2$ they correctly and completely solve instances (*i*), (*ii*) and (*iv*) in a coherent way, based on using existence and necessary first order optimality conditions, similar to parts (*a*) and (*b*) of Theorem 1, aided by considerations involving strict quasiconcavity of $u$ on $\mathbb{R}^2_{++}$: see the solutions to problems 1.2.1, 1.3.2 and 1.3.4 in [1, pp. 25-26, pp. 34-35]. Instance (*iii*) is not treated in [2] and instance (*v*) is solved in an *ad hoc* fashion; the former would not seem to be out of reach of the general method presented in [2], but the latter would seem to be. Thus, in the area of utility maximization the book [2], although not devoted to general microeconomic theory as such, managed to reach much further than the references mentioned above and it did so flawlessly.

For purposes going beyond computations, Definition 1.3.4 of [2] defines a *neoclassical preference* to be a continuous preference relation on $\mathbb{R}^\ell_+$ that has certain monotonicity properties. In terms of the representing utility function of such a preference relation, which exists by [10, Proposition 3.C.1], this definition comes down to the following two possibilities: either (1) $u$ is strongly increasing and strictly quasiconcave on $\mathbb{R}^\ell_+$ or (2) strongly increasing and strictly quasiconcave on $\mathbb{R}^\ell_{++}$ and $u(x) > u(x')$ for every $x \in \mathbb{R}^\ell_{++}$ and $x' \in \mathbb{R}^\ell_+ \backslash \mathbb{R}^\ell_{++}$. Here I use standard terminology from [8, 10]

– the one used in [2] is somewhat different. Observe that CES utility functions satisfy (1) but not (2) and that Cobb-Douglas utility functions satisfy (2) but not (1); thus, the above definition is inherently two-pronged. I observe that if possibility (2) obtains, then one has actually the following special property, which seems not to have been stated explicitly in [2]: everywhere on the boundary of $\mathbb{R}^\ell_+$ the function $u$ is equal to $u(0)$ (to see this, it suffices to compare any point on the boundary with $(t, t, \ldots, t)$ for sufficiently small $t > 0$ and use continuity and strict monotonicity of $u$ to finish the argument). I claim that both possibilities (1) and (2), and therefore the definition of a neoclassical preference relation as a whole, are subsumed by the more general notion of a strictly increasing utility function that is $(S)$-quasiconcave (incidentally, this continues to hold if in (1) and (2) "strongly increasing" is replaced by the less demanding "strictly increasing"). For possibility (1) this is immediately obvious. As for (2), by continuity of $u$ and strict quasiconcavity of $u$ on $\mathbb{R}^\ell_{++}$ it easily follows that $u$ is quasiconcave on $\mathbb{R}^\ell_+$, which is the closure of $\mathbb{R}^\ell_{++}$. To show property $(S)$, recall from the discussion of (2) in the preceding lines that if $x, x' \in \mathbb{R}^\ell_+$, $x \neq x'$, have $u(x) = u(x') > u(0)$ then this implies $x, x' \in \mathbb{R}^\ell_{++}$. Therefore, the fact that possibility (2) results in strict quasiconcavity of $u$ on $\mathbb{R}^\ell_{++}$ implies $u(\frac{1}{2}x + \frac{1}{2}x') > u(x) = u(x')$. As already said, this shows that utility functions that correspond to a neoclassical preference are $(S)$-quasiconcave. The converse is not true, even when the neoclassical preference definition is adapted so as to encompass strictly increasing utility functions: for instance, in Example 1$(iii)$ it was shown that in instance $(iv)$ the utility function $u$ is $(S)$-quasiconcave, but not strictly quasiconcave on $\mathbb{R}^2_+$; hence, for this $u$ possibility (1) is out of the question. On the other hand, (2) is also impossible because of $u(1, 0) \not< u(\frac{1}{2}, \frac{1}{2})$.

**Remark 3** $(i)$ Without monotonicity in Theorem 1, the budget-balancedness of the optimal solution can obviously not be maintained, but a fair part of Theorem 1 continues to hold when $u$ is nondecreasing and the details are as follows. In part $(a)$ the existence of at least one budget-balanced optimal solution is still guaranteed and part $(b)$ continues to hold as stated. Part $(c)$ of Theorem 1 remains meaningful by linking the additional possibility $x^* \in B\backslash B_0$ with the multiplier $\lambda = 0$. In that case the optimality of $x^*$ can be guaranteed if $u$, next to being nondecreasing, is concave on $\mathbb{R}^\ell_+$ (the example $\ell = 1$, $x^* = 1$, $u(x) := (x-1)^3$, $y = 2$ and $p = 1$ shows that mere quasiconcavity is insufficient in this situation). Finally, part $(d)$ is without significance: to have property $(S)$, a nondecreasing $u$ must be strictly increasing, a situation that is already covered by Theorem 1$(d)$ itself.

$(ii)$ Some utility functions have $\mathbb{R}^\ell_{++}$ as their natural domain of definition. Let $u : \mathbb{R}^\ell_{++} \to \mathbb{R}$ be such a function and suppose that $u$ is continuous and strictly increasing on $\mathbb{R}^\ell_{++}$, as well as differentiable on the open set $\Omega \subset \mathbb{R}^\ell_{++}$. To avoid trivialities, the new situation requires $y > 0$. Then parts $(b)$-$(d)$ of Theorems 1 continue to hold, naturally with quasiconcavity on $\mathbb{R}^\ell_+$ replaced by quasiconcavity on $\mathbb{R}^\ell_{++}$ and with property $(S)$ redefined as follows: for every pair $x, x' \in \mathbb{R}^\ell_{++}$ with $x \neq x'$ and $u(x) = u(x')$ one has $u(\frac{1}{2}x + \frac{1}{2}x') > u(x) = u(x')$. As shown by the example $\ell = 2$, $u(x_1, x_2) := x_1 + x_2$, part $(a)$ needs adjustment: an optimal solution exists under the extra condition that the set $C_v := \{x \in \mathbb{R}^\ell_{++} : u(x) \geq v\}$ is closed for every $v \in u(\mathbb{R}^\ell_{++})$. Namely, given $p$ and $y > 0$, fix any $\bar{x} \in \mathbb{R}^\ell_{++}$ with $p \cdot \bar{x} \leq y$ and set $\bar{v} := u(\bar{x})$. Then

$$\sup_{x \in \mathbb{R}^\ell_{++}, p \cdot x \leq y} u(x) = \sup_{x \in C_{\bar{v}}, p \cdot x \leq y} u(x)$$

and on the right side a continuous function is maximized over a nonempty compact set.

# 3 Testing for operational usefulness

Here I shall discuss the use of Theorem 1 as a means to meet my criterion of operational usefulness for utility maximization (in [3] similar applications were given for expenditure minimization). So my task is to derive complete solutions for the Marshallian demand function, using Theorem 1 in a coherent way, for each of the following utility functions:

$(i)$ $u(x) = \Pi_{i=1}^\ell x_i^{\alpha_i}$ with all $\alpha_i > 0$, (*ii*) $u(x) = (\sum_{i=1}^{\ell} x_i^{\rho})^{1/\rho}$ with $\rho \in (0,1)$,

(*iii*) $u(x) = \sum_{i=1}^{\ell} a_i x_i$, with all $a_i > 0$,

(*iv*) $u(x_1, x_2) = x_1^2(x_2 + 1)$,

(*v*) $u(x) = \min_{1 \le i \le \ell} b_i x_i$, with all $b_i > 0$.

These five functions are continuous and strictly increasing on $\mathbb{R}_+^{\ell}$. In instance (*v*) I suppose $\ell \ge 2$, so as to avoid overlap with instance (*iii*). My solution method, which is based on familiar reasoning in optimization theory [4], goes as follows. The starting point is that Theorem 1(*a*) guarantees that an optimal solution of (1) exists and is budget-balanced. It is useful to introduce the following term: an *optimality candidate* is a vector $x^* \in B_0 \cap \Omega$ that satisfies the first order necessary optimality condition (2). Then it follows from parts (*a*)-(*b*) of Theorem 1 that the optimal solution of (1) must be an optimality candidate, provided that it belongs to $\Omega$. Subsequently, if $u$ happens to be quasiconcave, then part (*c*) applies and all optimality candidates (if there are any) are indeed optimal solutions. If in addition $u$ has property ($S$), then the solution is of course complete (but only then, for I wish to determine *all* optimal solutions of (1)). In sum, if the utility function is ($S$)-quasiconcave, then an optimality candidate, when found, is immediately known to be the unique optimal solution of (1). If there does not exist an optimality candidate or if $u$ is not quasiconcave or fails to possess property ($S$), a careful look at the values that $u$ attains on the *remainder set* $B_0 \backslash \Omega$ is needed, and these values should be compared with the maximum value of all the optimality candidates already found (if any). For $\ell = 2$ the latter is easy, but for $\ell > 2$ it can be somewhat of a challenge: see my solution of instances (*iii*) and (*v*).

It must be mentioned that the alternative optimality conditions that I mentioned in Remark 1 can also be used, but only for instances such as (*iii*) and (*iv*), where the utility functions are differentiable on all of $\mathbb{R}^{\ell}$. As explained before, this gives then $\ell+1$ multipliers and, as is well-known from nonlinear programming, one should then work with $I(x^*) := \{i : i = 1, \ldots, \ell, x_i^* = 0\}$, the set of boundary-active indices of a bundle $x^* \in B_0$; cf. [4]. A good impression of this method can be obtained from the worked-out example of a linear technology in section 4.3 of [14] (the derivation there is nonrigorous because it is exclusively based on using first-order necessary conditions) . Another observation is that *ad hoc* methods to solve specific problems abound. For example, it is well-known that instance (*i*) can be solved by trivially eliminating the boundary of $\mathbb{R}_+^{\ell}$, after which one can apply the logarithmic transformation. However, such a transformation is already contained in Example 1(*i*), where it is part of a fairly systematic solution method.

**Solution (i)** I choose $\Omega = \mathbb{R}_{++}^{\ell}$. Let $u : \mathbb{R}_+^{\ell} \to \mathbb{R}$ be given by $u(x) := \Pi_{i=1}^{\ell} x_i^{\alpha_i}$, with all $\alpha_i > 0$. From Example 1(*i*) I know that $u$ is ($S$)-quasiconcave. So if I can find $x^* \in B_0 \cap \Omega$ that satisfies (2), then it must be the unique optimal solution. In search of such $x^*$, I combine (2) with $p \cdot x^* = y$ and verify concretely that the $x^*$ found belongs to $\Omega$. This is a simple algebraic task (incidentally, note that the possibility $\lambda = 0$ in (2) leads to $x^* \notin \Omega$). It yields $x_i^* = \alpha_i y/(\alpha p_i)$, $i = 1, \ldots, \ell$. This happens to be strictly positive, so the unique optimal solution has been found.

**Solution (ii)** I choose $\Omega = \mathbb{R}_{++}^{\ell}$. Let $u : \mathbb{R}_+^{\ell} \to \mathbb{R}$ be given by $u(x) := (\sum_{i=1}^{\ell} x_i^{\rho})^{1/\rho}$, with $\rho \in (0,1)$. From Example 1(*ii*) I know that $u$ is ($S$)-quasiconcave. So if I can find $x^* \in B_0 \cap \Omega$ that satisfies (2), then it must be the unique optimal solution. To find such $x^*$, I combine (2) with $p \cdot x^* = y$ and then check that their solution belongs to $\Omega$ (the possibility $\lambda = 0$ in (2) can be excluded, because it leads to nonsensical expressions). Again this is a simple algebraic task, which gives the strictly positive expression $x_i^* = p_i^{r-1} y/(p_1^r + \cdots + p_\ell^r)$, $i = 1, \ldots, \ell$, where $r := \rho/(\rho - 1)$. So the unique optimal solution has been determined.

**Solution (iii)** I shall solve (1) for the linear utility function $u(x) = \sum_{i=1}^{\ell} a_i x_i$, with all $a_i > 0$, choosing $\Omega = \mathbb{R}_{++}^{\ell}$, as before. Let $a$ be the $\ell$-vector with components $a_i$. Exactly one of the following can occur. Case 1: $a$ is a scalar multiple of $p$ (say $a = \mu p$ for some $\mu \in \mathbb{R}$, and then $\mu > 0$ of course) or case 2: $a$ is not a scalar multiple of $p$.

*Case 1:* By Theorem 1(*b*), for any optimal $x^* \in \Omega := \mathbb{R}_{++}^{\ell}$ there exists $\lambda \ge 0$ such that $a = \lambda p$. In the present case this was already true (take $\lambda = \mu > 0$), so every $x^* \in B_0 \cap \Omega$ is an optimality candidate. Next, $u$ is quasiconcave, so every optimality candidate is also an optimal solution. Again this offers no news, because all optimality candidates $x^*$ satisfy $a \cdot x^* = \mu p \cdot x^* = \mu y$ and because $a \cdot x = \mu p \cdot x \leq \mu y$ for all $x \in B$. It remains to inspect the remainder set $B_0 \backslash \Omega$ (here that is the set of all $x^*$ in $B_0$ with at least one coordinate equal to zero). The value $u(x^*)$ of any $x^* \in B_0 \backslash \Omega$ is $a \cdot x^* = \mu p \cdot x^* = \mu y$, which is the same value as found before. Hence, I conclude that in case 1 the set of all optimal solutions is $B_0$, that is to say the union of $B_0 \cap \Omega$ and the remainder set $B_0 \backslash \Omega$.

*Case 2:* This time $a = \lambda p$ is incompatible with the situation of the present case. So there are no optimality candidates at all, which means that the optimal solution (which is known to exist) must belong to the remainder set $B_0 \backslash \Omega$. Now for every $x$ in $B$ I have $a \cdot x = \sum_i a_i p_i^{-1} p_i x_i \leq \alpha p \cdot x \leq \alpha y$, where $\alpha := \max_i a_i/p_i$. Let $I$ be the set of those indices $i$ for which $a_i/p_i = \alpha$. Then $I$ is nonempty and it is a strict subset of $\{1, \ldots, \ell\}$ (or else I would find $a = \alpha p$, which cannot be true in the present case). Now any vector $\frac{y}{p_i} e_i$, $i \in I$, belongs to $B_0$ and achieves $a \cdot \frac{y}{p_i} e_i = \alpha y$. Here $e_i$ denotes the $i$-th unit vector in $\mathbb{R}^\ell$. So I conclude $\sup_{x \in B} a \cdot x = \alpha y$. I claim that the set of optimal solutions is the intersection $B_0 \cap \cap_{i \notin I} \{x \in \mathbb{R}^\ell_+ : x_i = 0\}$, which is a face of $B_0$. First, let $x^*$ belong to this intersection. Then $a \cdot x^* = \sum_{i \in I} a_i x_i^* = \sum_{i \in I} a_i p_i^{-1} p_i x_i^* = \alpha p \cdot x^* = \alpha y$, so $x^*$ is optimal. Next, let $x^*$ be optimal. Then $x^* \in B_0$ by Theorem 1$(a)$ and $\alpha y = a \cdot x^*$ is given, so $\alpha y = \alpha \sum_{i \in I} p_i x_i^* + \sum_{i \notin I} a_i x_i^*$. By $p \cdot x^* = y$ this implies $\sum_{i \notin I} (a_i - \alpha p_i) x_i^* = 0$, so it follows that $x_i^* = 0$ for every $i \notin I$ (note that $a_i - \alpha p_i < 0$ for each $i \notin I$). This proves the desired characterization of the set of optimal solutions in case 2.

**Solution (iv)** I solve (1) for the utility function $u(x_1, x_2) = x_1^2(x_2 + 1)$, taking $\Omega = \mathbb{R}^2_{++}$. From Example 1$(iii)$ I already know that $u$ is $(S)$-quasiconcave. So if I can find $x^* \in B_0 \cap \Omega$ that satisfies (2), then it must be the unique optimal solution. I solve $p \cdot x^* = y$ and (2), the latter amounting to $x_2^* = p_1(2p_2)^{-1} x_1^* - 1$, and then pick the solution, if any, that belongs to $\Omega$. From the former two equations I also obtain that $\lambda = 0$ if and only if $(x_1^*, x_2^*) = (0, y/p_2)$, a vector that does not belong to $\Omega$. So I can proceed with $\lambda > 0$. I solve the two equations, which gives $x_1^* = \frac{2}{3}(y + p_2)/p_1$ and $x_2^* = -1 + \frac{1}{3}(y + p_2)/p_2$. Because $x^* \in \Omega$ is needed for $x^*$ to be an optimality candidate, I distinguish between the following two cases (note that $y > 2p_2$ is equivalent to $x^* \in \Omega$):

*Case 1:* $y > 2p_2$. Then $x^* = (\frac{2}{3}(y + p_2)/p_1, -1 + \frac{1}{3}(y + p_2)/p_2)$ belongs to $\Omega$. Hence it is an optimality candidate and by the previous argument it must also be the unique optimal optimal solution.

*Case 2:* $y \leq 2p_2$. In this case there is not any optimality candidate. Therefore, I know that the best $u$-value over the remainder set $B_0 \backslash \Omega$ is the optimal value. The set $B_0 \backslash \Omega$ contains only two vectors, the corner points $(y/p_1, 0)$ and $(0, y/p_2)$. Of these, the former one gives the highest $u$-value. Combining cases 1 and 2, I conclude that the Marshallian demand function is given by

$$(x_1^*, x_2^*) = \begin{cases} (2(y + p_2)/3p_1, (y - 2p_2)/3p_2) & \text{if } y > 2p_2, \\ (y/p_1, 0) & \text{if } y \leq 2p_2, \end{cases}$$

and it is illuminating to illustrate this graphically.

**Solution (v)** I solve (1) for $u(x) = \min_{1 \leq i \leq \ell} b_i x_i$, with all $b_i > 0$ and I choose $\Omega$ to be the open set of all $x \in \mathbb{R}^\ell_{++}$ such that $b_i x_i \neq b_j x_j$ for any $i \neq j$. Then $u$ is locally of the form $u(x) = b_k x_k$ near each point $x^*$ of $\Omega$ (the index $k$ being unique to $x^*$), so there are no optimality candidates at all, because in such $x^*$ condition (2) amounts to $b_k e_k = \lambda p$, which does not have a solution (recall that $\ell \geq 2$). Here $e_k$ again denotes the $k$-th unit vector. It follows that the optimal solution, known to exist, must belong to the remainder set $B_0 \backslash \Omega$. For any $x$ in the remainder set $B_0 \backslash \Omega$ either of the following two cases can occur: case 1: $x \in \mathbb{R}^\ell_{++}$ and there is a tie in the form of an equality $b_i x_i = b_j x_j$ for some $i \neq j$ or case 2: some coordinate of $x$ is equal to zero. In case 2 $u(x)$ equals zero, which means the certain sub-optimality of $x$. So I concentrate on those vectors in $B_0 \cap \mathbb{R}^\ell_{++}$ that have ties. For any $x$ in $B$ I have $x_i \geq u(x)/b_i$ for all $i$, giving $y \geq \sum_i p_i x_i \geq u(x) \sum_i p_i/b_i$, i.e., $u(x) \leq y/\beta$, with $\beta := \sum_i p_i/b_i > 0$. In fact, the same reasoning shows that $u(x) < y/\beta$ whenever the set of all indices $i$ for which $b_i x_i > u(x)$ is nonempty. So $u(x^*) = y/\beta$ for $x^* \in B$, which amounts to $x^*$ being optimal in view of the previous lines, requires $b_i x_i^* = b_j x_j^* = u(x^*)$ for all $i$ and $j$. Clearly, there is precisely one such $x^*$ and it is given by $x_i^* = b_i^{-1} y/\beta$, $i = 1, \ldots, n$; so this is the optimal solution.

**Acknowledgments.** I am indebted to Jean-Pierre Crouzeix (Université Blaise Pascal, France) for the reported observations in connection with reference [6] and to Roko Aliprantis (Purdue University, USA) for calling my attention to his books [1, 2] and for helpful remarks. I thank Alexander Gnedin (University of Utrecht) for helpful criticism and encouragement.

## References

[1] Aliprantis, C.D. *Problems in Equilibrium Theory.* Springer-Verlag, Berlin, 1995.

[2] Aliprantis, C.D., Brown, D.J. and Burkinshaw, O. *Existence and Optimality of Competitive Equilibria.* Springer-Verlag, Berlin, 1989.

[3] Balder, E.J. A mathematical introduction to microeconomics, lecture notes WISB172, Department of Mathematics, University of Utrecht, 2008.

[4] Bazaraa, M.S., Sherali, H.D. and Shetty, C.M. *Nonlinear Programming* (2nd ed.), Wiley, New York, 1993.

[5] Besanko, D. and Braeutigam, D.R. *Microeconomics* (2nd. ed.), Wiley, New York, 2005.

[6] Crouzeix, J.-P. Conditions for convexity of quasiconvex functions, Mathematics of Operations Research 5 (1980), 120-125.

[7] Edwards, C.H. *Advanced Calculus of Several Variables*, Academic Press, New York, 1973.

[8] Jehle, G.A. and Reny, P.J. *Advanced Microeconomic Theory* (2nd edition), Wiley, New York, 2001.

[9] Luenberger, D.G. *Microeconomic Theory*, McGraw-Hill, New York, 1995.

[10] Mas-Colell, A., Whinston, M.D. and Green, J.R. *Microeconomic Theory*, Oxford University Press, New York, 1995.

[11] Simon, C.P. and Blume, L. *Mathematics for Economists*, Norton, New York, 1993.

[12] Sydsæter, K., Hammond, P., Seierstad, A. and Strøm, A. *Further Mathematics for Economic Analysis*, Prentice Hall, Harlow, England, 2005.

[13] Takayama, A. *Mathematical Economics* (2nd edition), Cambridge University Press, Cambridge, 1996.

[14] Varian, H.R. *Microeconomic Analysis* (3rd edition), Norton, New York, 1992.