WORKING PAPER SERIES

On the Recursive Saddle Point Method

Matthias Messner and Nicola Pavoni

Working Paper n. 255

February 2004

On the Recursive Saddle Point Method [*]

Matthias Messner[†]   Nicola Pavoni[‡]
Bocconi University and IGIER   University College London

February 3, 2004

Abstract

In this paper a simple dynamic optimization problem is solved with the help of
the recursive saddle point method developed by Marcet and Marimon (1999). Ac-
cording to Marcet and Marimon, their technique should yield a full characterization
of the set of solutions for this problem. We show though, that while their method
allows us to calculate the true value of the optimization program, not all solutions
which it admits are correct. Indeed, some of the policies which it generates as
solutions to our problem, are either suboptimal or do not even satisfy feasibility.
We identify the reasons underlying this failure and discuss its implications for the
numerous existing applications.

Keywords: Recursive saddle point, recursive contracts, dynamic programming.
JEL code: C61, C63.

1 Introduction

Marcet and Marimon (1999) (henceforth MM) develop an elegant and flexible recursive saddle point method, suitable for solving a wide class of dynamic optimization problems. Because of its tractability and computational advantages, throughout the last years many authors have started applying this approach.[1] The list of papers in which it is used includes, among many others, Attanasio and Rios-Rull(2001), Ayagari, Marcet, Sargent and Seppälä (2002), Cooley, Marimon and Quadrini (2001), Friedman (1998), Kehoe and Perri (2002), Khan, King and Wolman (2000), Klein and Rios-Rull (2002), Marcet and Marimon (1992), Seppälä (2002), and Siu (2002).

In this paper we apply the MM method to a simple concave dynamic optimization problem which falls into the class of problems which MM consider in their paper. In this example a principal has to share a constant stream of endowments with an agent in such a way that his own utility is maximized subject to the constraint that the agent's utility never falls below a certain threshold. We show that while the MM technique allows one to calculate the true value of the problem, the set of solutions which it admits does not coincide with the true set of solutions. In particular, in our example the MM approach allows both for solutions which are feasible but suboptimal and solutions which even violate feasibility.

The method developed by MM builds on the fact that the solution of an optimization problem can often be obtained by finding the saddle points of its associated Lagrangean. Roughly speaking, the idea underlying the MM approach is to use recursive techniques to calculate these saddle points. That is, MM aim to show that by using appropriate summaries of the Lagrange multipliers as state variables (MM refer to these state variables as *co-states*) a sequential saddle point problem can be transformed into a recursive one which gives rise to exactly the same set of solutions.[2] In particular, they argue that the equivalence between sequential and recursive saddle points does not require any concavity assumptions on the underlying optimization problem. They conclude therefore that, just as the standard Lagrangean approach, their method provides a full characterization of the set of solutions of concave optimization programs, and yields sufficient conditions for solutions of nonconcave problems.[3]

---

[1] One of the key advantages of the MM method with respect to dynamic programming is that the state space is not itself endogenous but is given exogenously. This can simplify the numerical the numerical analysis substantially as costly preliminary computations which the Bellman approach might imply can be avoided (see, for example, Abraham and Pavoni (2003) and Chang (1998)).

[2] Hence, they essentially extend Bellman's *Principle of Optimality* for dynamic optimization problems to (a class of) dynamic saddle point problems.

[3] In their own words: 'Concavity is no more necessary for our approach than for the classical Lagrangean method.' (MM page 3).

Since our example is concave, equivalence between saddle points of the Lagrangean and solutions of the optimization problem indeed holds. The fact that the MM method nonetheless allows for wrong solutions therefore proves that the set of sequential saddle points and the set of their recursive counterparts do not always coincide under the conditions imposed by MM. For our problem it turns out, that the latter is a strict superset of the first one. Hence, at least in this case the technique of MM yields only necessary but not sufficient conditions for a solution.

In our discussion we will argue that the MM method might fail to yield a correct solution whenever the problem under consideration is not strictly concave. This highlights the importance of our results: Concavity (and hence all the more so strict concavity) is a rather restrictive condition as many interesting economic problems which include incentive constraints are very often nonconcave. In fact, most of the models analyzed in the articles mentioned above study problems which fall into two main categories, Ramsey taxation and default with capital accumulation, which are known to have typically a nonconcave structure.[4] An exception to this point is Attanasio and Rios-Rull (2001), who study risk sharing with default without capital accumulation. It seems that the popularity of the MM approach derives mainly from the computational advantages it provides in numerical analyses (see footnote 1). Our findings (the lack of sufficiency of the optimality conditions which the method yields) show though that its use for such purposes is rather inappropriate.

An interesting interpretation of the limitations of the MM approach can be obtained by comparing it with classical dynamic programming. This comparison reveals that the MM co-states do not allow for a sufficiently 'sharp' description of the true state of the optimization problem. In a sense which we will make more precise in this paper, in our example to each of the MM co-states along the optimal path there corresponds a whole interval of values of the *primal* state variable used in the Bellman approach. Therefore, fixing a co-state does not pin down the correct continuation path which, as we will see, in effect amounts to a relaxation of the true feasibility and optimality conditions.

We proceed in the following way. In the next section we introduce our problem and describe its true set of solutions. In Section 3 we go on to characterize the set of solutions obtained by using the MM method. In Section 4 we discuss the results and the reasons underlying the failure of the recursive saddle point technique. The last

---

[4]Ramsey taxation problems are known to be in general nonconcave since Lucas and Stokey (1983) (see page 62). In the case of default with capital accumulation Cooley et al. (2001) themselves (in footnote 10) point out that the problem is nonconcave since the endogeneity of default value creates nonconvexities in the incentive feasibility set.

section concludes.

2   The Problem

Consider the following problem:

$$\sup_{\{a_t\}} \quad \sum_{t=0}^{\infty} \beta^t \left(y - a_t\right) \tag{1a}$$

$$s.t. \quad a_t \in [0, \bar{a}], \quad \forall t \geq 0 \tag{1b}$$

$$\sum_{n=0}^{\infty} \beta^n u(a_{t+n}) \geq \frac{b}{1 - \beta}, \quad \forall t \geq 0, \tag{1c}$$

where $y > \bar{a} > b > 0$ and $0 < \beta < 1$.[5] Throughout the main part of the paper we will furthermore assume that $u(a) = a$. Only in the appendix we will consider the case of a non-linear constraint.

For the sake of concreteness we will interpret this problem as one where a sequence of endowments $(y, y, ...)$ has to be divided between two agents in such a way, that the discounted value of the consumption stream of the first one (1a) is maximized, subject to a technological feasibility condition (1b) and the constraint that the discounted value of the consumption stream of the second household never falls below $b/(1 - \beta)$ (1c). Throughout the paper we will refer to the first consumer also as planner or principal and to the second one as agent.[6]

Given the linearity of the utility functions of both individuals and the simple form of the technological feasibility constraint, the problem is rather trivial and no specific techniques are required to characterize the set of its solutions. The set of optimizing sequences, which we denote by A^*, is given by all sequences $\{a_t\}$ which satisfy the following conditions:

$$a_t \in [0, \bar{a}] \quad \forall t, \tag{2a}$$

$$\sum_{t=0}^{\infty} \beta^t a_t = \frac{b}{1 - \beta}, \text{ and} \tag{2b}$$

$$b \sum_{n=0}^{t-1} \beta^n - \sum_{n=0}^{t-1} \beta^n a_n \geq \beta^t (a_t - b) \quad \forall t \geq 0. \tag{2c}$$

---

[5]Notice that throughout we stick to the notation of MM in order to facilitate the comparison.

[6]In the terminology of optimal contracting one can interpret our problem also as one where a planner maximizes his discounted returns subject to the constraint that the agent should have no incentive to 'default'(condition (1c)).

The second condition simply requires that the agent receives a discounted utility exactly equal to $b/(1-\beta)$ in period zero. The third condition instead says that optimal consumption plans of the agent have to be 'backloaded': Transfers to the agent can exceed b only by the accumulated (and appropriately discounted) amount by which the payments in the periods up to t have fallen short of b.[7] Notice that the value generated by any sequence in A^* is $(y-b)/(1-\beta)$.

Since the programm (1a)-(1c) is concave, each element of A^* corresponds to a solution of the saddle point problem

$$\inf_{\{\lambda_t\}\in\mathbb{R}_+^\infty} \sup_{\{a_t\}\in[0,\bar{a}]^\infty} \sum_{t=0}^\infty \beta^t \left[y - a_t + \lambda_t \left(\sum_{n=0}^\infty \beta^n a_{t+n} - b \right) \right]. \tag{3}$$

In particular one can easily verify that $\{a_t\}$ belongs to A^* if and only if $((1,0,0,...),\{a_t\})$ is a solution of (3).

3 Characterizing the set of solutions using the MM approach

In this section we characterize the set of solutions to our problem with the help of the method developed by MM.[8] Essentially, their approach consists in transforming the saddle point problem associated to the original sequential maximization problem into a recursive form. In our case, this yields the following recursive saddle point functional equation:[9]

$$
\begin{aligned}
W(\mu^0,\mu^1) &= \sup_{a\in[0,\bar{a}]} \inf_{\substack{\gamma^0,\gamma^1\geq0,\\ \mu^{0'},\mu^{1'}\geq0}} \mu^0\left(y-a\right) + \gamma^0\left(y-a-R\right) + \mu^1 a + \gamma^1(a - \frac{b}{1-\beta}) + \beta W(\mu^{0'},\mu^{1'})\\
&\quad s.t. \quad \mu^{0'} = \mu^0 + \gamma^0\\
&\qquad\qquad \mu^{1'} = \mu^1 + \gamma^1
\end{aligned}
\tag{4}
$$

---

[7]Combining this last condition with (2b) yields again the no-default constraint (1c).

[8]Notice that our problem is indeed contained in the class of problems addressed by MM. Properties **A1** and **A2** (see page 19 of MM) of MM are trivially satisfied, as our problem is deterministic and in its original sequential form does not include a state variable. Also, since both the objective function as well as the constraint are linear, they are not only continuous, but also quasiconcave. The boundedness requirements in **A3** and **A4** are not an issue here since we can always define both the return of the planner and the utility of the agent on $[0,y]$. Moreover, the space of the sequences $\{a_t\}$ which satisfy technical feasibility is a convex set as it is the Cartesian product of convex sets. Finally, any constant transfer stream $\{a,a,...\}$ with $b < a < \bar{a}$ satisfies the interiority condition **A5**.

[9]For more details on how to derive this recursive saddle point functional equation from the sequential saddle point problem see Section 2 of MM.

As in MM we have introduced the constant R, which can be any number which bounds the planner's per period payoff from below. Since in our case the principal's consumption can never be negative ($\bar{a} < y$), we can set $R = 0$.

Suppose W^* solves problem (4) and ϕ_a and ϕ_γ are the corresponding policy correspondences. MM argue that any sequence of transfers generated with these policy correspondences when starting with the initial condition $\mu^0_{-1} = 1$ and $\mu^1_{-1} = 0$ belongs to A^* and vice versa.

Proposition 1. *The function $W^* : \mathbb{R}^2 \to \mathbb{R}$ defined by*

$$W^*(\mu^0, \mu^1) = \begin{cases} \mu^0 \dfrac{y - b}{1 - \beta} + \mu^1 \dfrac{b}{1 - \beta} & \text{if } \mu^0 \geq \mu^1 \\ \mu^0 \dfrac{y - \bar{a}}{1 - \beta} + \mu^1 \dfrac{\bar{a}}{1 - \beta} & \text{if } \mu^0 < \mu^1 \end{cases}$$

solves the functional equation (4).[10] *The corresponding policy correspondences ϕ_a : $\mathbb{R}^2_+ \to \mathcal{P}([0, \bar{a}])$ and $\phi_\gamma : \mathbb{R}^2_+ \to \mathbb{R}^2_+$ satisfy*

$$\phi_a(\mu^0, \mu^1) = \begin{cases} \bar{a} & \text{if } \mu^0 < \mu^1 \\ [\max\{0, (b - \beta\bar{a})/(1 - \beta)\}, \bar{a}] & \text{if } \mu^0 = \mu^1 \\ [\max\{0, (b - \beta\bar{a})/(1 - \beta)\}, b] & \text{if } \mu_0 > \mu_1 \end{cases}$$

and

$$\phi_\gamma(\mu^0, \mu^1) = \begin{cases} (0, 0) & \text{if } \mu^0 < \mu^1 \\ (0, \mu^0 - \mu^1) & \text{if } \mu^0 \geq \mu^1. \end{cases}$$

Proof. We simply have to show that for the given value function W^*, the policies ϕ_a and ϕ_γ solve the saddle point problem in (4) and that plugging these solutions back into the saddle point problem returns again W^*.

Given W^* the saddle point problem can by rewritten with the help of the constraints as:

$$\sup_{a \in [0, \bar{a}]} \inf_{\gamma^0, \gamma^1 \geq 0} \mu^0 (y - a) + \gamma^0 (y - a) + \mu^1 a + \gamma^1 (a - \frac{b}{1 - \beta}) + \beta W^*(\mu^0 + \gamma^0, \mu^1 + \gamma^1) \quad (5)$$

Observe, that (5) is strictly increasing in γ^0, regardless of the values of γ^1 and a and also independently of whether $\mu^0 \geq \mu^1$ or $\mu^0 < \mu^1$ (remember, that $\bar{a} < y$ by assumption). Therefore, as claimed, the unique optimal value for γ^0 is zero.

Next, suppose $\mu^0 < \mu^1$. We have to show that in this case the unique solution of the saddle point problem in (5) is given by $\gamma^1 = 0$ and $a = \bar{a}$. In order to do so, notice first, that for all $\gamma^1 \geq 0$ the objective function is strictly increasing in a. Therefore,

---

[10]Notice, that, consistently with the results of MM, the value function W^* is both continuous and homogeneous of degree one (see Proposition 4 of MM).

setting $a = \bar{a}$ is the unique optimal choice. Conversely, if $a = \bar{a}$, the slope of (5) in γ^1 is

$$\bar{a} - \frac{b}{1-\beta} + \beta \frac{\bar{a}}{1-\beta} = \frac{\bar{a} - b}{1-\beta}.$$

Since $\bar{a} > b$, it follows that the unique minimizing value for γ^1 is zero.

Finally, consider the case $\mu^0 \geq \mu^1$. We first argue, that the only candidate for the choice of γ_1 is $\mu_0 - \mu_1$. The slope of (5) in γ_1 is given by[11]

$$a - \frac{b}{1-\beta} + \beta \begin{cases} \dfrac{b}{1-\beta} & \text{if } \gamma^1 \leq \mu^0 - \mu^1 \\[2mm] \dfrac{\bar{a}}{1-\beta} & \text{if } \gamma^1 \geq \mu^0 - \mu^1. \end{cases}$$

Hence, in order for a finite minimizer to exist (for γ^1), a must be chosen greater than $(b - \beta\bar{a})/(1-\beta)$. This, together with technological feasibility, requires that $a \in [\max\{0, (b - \beta\bar{a})/(1-\beta)\}, \bar{a}]$. Values of γ_1 strictly larger than $\mu^0 - \mu^1$ require $a = (b - \beta\bar{a})/(1-\beta)$ (only for this value of a (5) is constant in γ^1). But for all $\gamma^1 > \mu^0 - \mu^1$ (5) is strictly increasing in a and hence a would have to be set equal to $\bar{a}$; a contradiction. On the other hand, for all values of γ^1 from the interval $[0, \mu^0 - \mu^1)$ the objective function is strictly decreasing in a, implying an optimal value of a equal to zero. For $a = 0$ though, the Lagrange multiplier would have to be at least as large as $\mu^0 - \mu^1$ as the slope of (5) over this range is $-b$ for $a = 0$. This again leads to a contradiction. It follows therefore that $\gamma^1 = \mu^0 - \mu^1$ is the unique candidate for the saddle.

If γ^1 equals $\mu^0 - \mu^1$, a cancels from (5) and hence, all technically feasible values would be maximizers. It remains to be determined for which subset of $[0, \bar{a}]$ $\gamma^1 = \mu^0 - \mu^1$ is a minimizer. We have already seen in the previous paragraph, that a must satisfy, $a \geq (b - \beta\bar{a})/(1-\beta)$ as this guarantees that the objective function is non-decreasing for $\gamma^1 > \mu^0 - \mu^1$. This is the only relevant condition if $\mu^0 - \mu^1 = 0$. Hence, in that case any a in $[\max\{0, (b - \beta\bar{a})/(1-\beta)\}, \bar{a}]$ is admissible. If instead, $\mu^0 - \mu^1 > 0$ then we also have to impose a condition that assures that (5) is non-increasing for $\gamma^1 < \mu^0 - \mu^1$ (this is required in order to avoid that $\gamma^1 = 0$, which we know cannot be part of a saddle point). As can be seen by inspection of the slopes of the objective function, this condition will be satisfied whenever $a \leq b$.

This completes the part of the proof regarding ϕ_a and ϕ_γ. It is a simple algebraic exercise to show that plugging the corresponding values of the policies back into (5) returns the correct expression for W^*. $\qquad\square$

Remember, that the initial values for μ^0 and μ^1 are 1 and 0, respectively. These initial conditions together with the above stated policies imply that in the first period

---

[11]The slopes in $\gamma^1 = \mu^0 - \mu^1$ given in the following expression should be interpreted as the righthand and lefthand derivatives respectively.

a has to be chosen from $[\max\{0, (b - \beta\bar{a})/\ (1-\beta)\}, b]$, while γ_1 is to be set equal to $\mu^0 - \mu^1 = 1$. From the second period on the state will remain constant, since according to ϕ_γ, γ^1 must be zero whenever $\mu^0 = \mu^1$. Consequently, we also have a fixed set of admissible choices for the control a from the second period onwards, which is equal to $[\max\{0, (b - \beta\bar{a})/(1-\beta)\}, \bar{a}]$.

Summarizing, the set of transfer sequences $\{a_t\}$, that satisfy the optimality conditions of the MM approach, A^{MM}, is characterized by

$$a_t \in [0, \bar{a}] \qquad \forall\, t, \tag{6a}$$

$$a_0 \in [(b - \beta\bar{a})/(1-\beta), b] \quad \text{and} \tag{6b}$$

$$a_t \in [(b - \beta\bar{a})/(1-\beta), \bar{a}] \quad \forall\, t \geq 1. \tag{6c}$$

A first inspection of these conditions immediately reveals that they do not specify any intertemporal link between the controls. This is in clear contrast to the defining conditions of A^*. Both (2b), which requires the agent to get a utility of exactly $b/(1-\beta)$ and (2c), which restricts possible solutions to backloaded paths, define intertemporal relations between the controls.

In order to highlight the consequences of this divergence between (2a)-(2c) and (6a)-(6c), suppose that $b < \beta\bar{a}$. According to (6a)-(6c), in this case the sequence of zero-transfers, $a_t = 0 \; \forall\, t$, also belongs to A^{MM}. But of course, such a sequence is not (incentive) feasible in the sequential problem and hence does not belong to A^*, as it implies a zero lifetime utility for the agent in each period, while by hypothesis he should receive at least $b/(1-\beta)$ (see (1c)). So we have to conclude that the recursive saddle point approach allows for 'wrong' solutions.

The zero sequence, is of course only one of many wrong solutions. In fact, with the sequences contained in A^{MM} one can generate any payoff for the planner that lies between $y/(1-\beta) - b - \beta\bar{a}/(1-\beta)$ and y. Hence, some of the MM-solutions yield the planner a lower payoff than the truely optimal transfer schemes. While not being optimal some of those sequences satisfy feasibility (take the sequence $(b, \bar{a}, \bar{a}, ...)$).

Notice finally, that all sequences that satisfy (2a)-(2c) also satisfy (6a)-(6c).[12] In other words, the conditions (6a)-(6c) derived with the help of the MM method are only necessary but not sufficient conditions for the true set of solutions.

---

[12]Technical feasibility is obviously satisfied. Moreover, setting $t = 0$ in (2c) gives $a_0 \leq b$ as required in (6b). From the same condition follows that one must have $a_t \geq \max\{0, (b - \beta\bar{a})/(1-\beta)\}$ for all $t \geq 0$: For any a smaller than $(b - \beta\bar{a})/(1-\beta)$, not even continuing with the maximal transfers $\bar{a}$ in all future periods would allow to satisfy the no-default condition (2c).

4 Discussion

4.1 Why suboptimal and/or unfeasible solutions?

Why is the MM-approach not able to isolate the truly optimal solutions? In order to understand the reasons underlying this failure, let us consider the conditions (6a)-(6c), that characterize A^{MM} in some more detail. They tell us, that in every period we must have $a \geq (b - \beta\bar{a})/(1 - \beta)$, or equivalently $a + \beta\frac{\bar{a}}{1 - \beta} \geq \frac{b}{1 - \beta}$. The left hand side of this last expression is nothing else but the agents utility, if he is given a today and $\bar{a}$ in every period from tomorrow onwards. We can therefore interpret the whole condition as the *recursive form* of the no-default condition in (1c): a is an admissible choice today as long as there is at least *one* (technically) feasible future stream of transfers, which together with a guarantees the agent a payoff (from today onwards) no lower than $b/(1 - \beta)$.

This condition per se makes perfectly sense. The problem is simply that the MM-approach fails to 'enforce' it. As we have already pointed out in the previous section, there is nothing in the MM-optimality conditions which links the choices of the controls across different periods. Instead, the set of admissible choices is constant throughout time,[13] and so the method is not able to guarantee that a low transfer payment in the current period will be followed by sufficiently high payments in the future. The zero-transfer example of the previous section demonstrates this most clearly. In each period t, the payment $a_t = 0$ is 'acceptable' since the overall payoff equal to $b/(1 - \beta)$ would be guaranteed, for example, by the continuation path $(b/\beta, b/\beta, b/\beta, ...)$.

As we have already remarked in the introduction, it is useful to compare the MM-method with classical dynamic programming in order to understand its failure better.[14]

As in many economic applications the appropriate state variable which allows us to set up the optimization problem (1a)-(1c) in the form of a Bellman equation is the agent's continuation utility from period t onward $U_t = \sum_{n=0}^{\infty} \beta^t a_{t+n}$. It satisfies the following (implicit) law of motion:

$$U_t = a_t + \beta U_{t+1}. \tag{7}$$

It is well known that the value of our optimization problem (1a)-(1c) is given by

---

[13]Only in the initial period the set of admissible choices is different from $[\max\{0, (b-\beta\bar{a})/(1-\beta), \bar{a}\}]$.

[14]For a more detailed discussion of what follows see Stokey et. al. (1989).

$V(b/(1-\beta))$, where the function V solves the following functional equation

$$V(U) = \sup_{a\in[0,\bar{a}],U'\in\left[\frac{b}{1-\beta},\frac{\bar{a}}{1-\beta}\right]} y - a + \beta V(U') \tag{8}$$

$$\text{s.t.} \quad a + \beta U' \geq \frac{b}{1-\beta} \tag{9}$$

$$U = a + \beta U'. \tag{10}$$

One can easily verify that the set of solutions generated by this method coincides with A^*.

Notice, that the law of motion of the state variable U enters as a constraint in the Bellman equation. Hence, with every admissible value for the control a today, one chooses also the *unique discounted value* of the sequence of controls to be followed in the future, U'. Since U' in turn will define the constraint for the next period, the consistency of the choice throughout time is guaranteed, that is, if in the current period (a, U') are chosen, U' is indeed the utility that the agent will receive in the future periods.

The distinguishing feature of the MM-approach as opposed to classical dynamic programming is that it does not set up the problem in terms of the planners objective. Instead it transforms the original optimization program into a *recursive welfare maximization problem*, where the welfare weights attached to the principal (μ^0) and to the agent (μ^1) serve as state variables. Hence, in the MM method, the continuation utility of the agent is not chosen directly but is controlled by fixing corresponding welfare weights. In this sense, the MM-states provide an 'indirect' or 'dual' description (as opposed to the 'primal' states used in dynamic programming) of the state of the system.

This 'indirect' approach works fine as long as we can associate to each dual state a unique primal state, but it generates wrong solutions as soon as a one-to-one relationship between the two types of states fails to exist. This is exactly what happens in our example: Due to our linearity assumptions, the frontier of the set of technically feasible continuation utilities is decreasing at the constant slope of -1. Therefore, any point on this frontier corresponds to a solution of the welfare maximization problem if the welfare weights of both agents are the same.

The consequences of this are very clear: If a single pair of welfare weights corresponds to all possible divisions of the joint welfare of the two individuals from the current period onwards, this simply means that the MM approach does not allow for a sufficiently 'sharp' description of the state. Fixing a dual state does not imply a specific promise about how to divide the pie in the future. Instead, all possible divisions which correspond to the given pair of welfare weights might be followed.

It is interesting to notice, that choosing a solution in $A^{MM} - A^*$ does not imply, that we calculate a wrong value for the planner. Evaluating W^* at the point $(1,0)$ gives $(y-b)/(1-\beta)$, which is indeed the value of the original optimization problem as we have already seen in Section 2. The intuition for this is that in a recursive welfare maximization problem, the value in period zero does not depend on how the welfare is distributed between the two individuals from the second period onwards. All what matters is that in every period the μ-weighted sum of the two agents welfare is maximized. For $(\mu^0, \mu^1) = (1,1)$ this condition is satisfied by all points on the linear utility possibility frontier and so shifting welfare between the two agents in any non-initial period has no influence on the (period 0) value of W^* at $(1,0)$.

4.2 Recursive vs. sequential saddle points and the role of concavity

MM link the solutions obtained with their method only in an indirect way with the true set of solutions of the underlying optimization problem. Their strategy can be described as follows: After defining a class of optimization problems they characterize the relationship between the solution sets of those problems and the saddle points of the corresponding Lagrangeans. Using standard arguments they show that the Lagrangean method always yields sufficient conditions for a solution which are also necessary in case the problem under consideration is concave.

In a second step they then go on to develop their main result. They show that any sequential saddle point problem associated to an optimization problem from their class, can be transformed into a recursive saddle point problem which under rather weak conditions gives rise to the same set of solutions as the former. In fact, they argue that whenever the underlying optimization problem is concave no further conditions are required. For non-concave problems, recursively calculated saddle points solve the corresponding sequential saddle point problem only if the sequence of the recursively generated Lagrange multipliers satisfies a boundedness condition (while again for the other direction no additional conditions are necessary).

Combining the two results, MM conclude that their method not only yields a full characterization of the set of solutions of a concave optimization problem but that it also gives sufficient conditions (though not necessary ones) for non-concave problems (subject to the above mentioned boundedness condition).

We have already pointed out in Section 2, that for any solution of our concave problem there is a corresponding saddle point of the associated Lagrangean and vice versa. The fact that the MM method yields wrong solutions for our problem tells us therefore, that concavity alone cannot be a sufficient condition for equivalence between sequentially and recursively calculated saddle points.
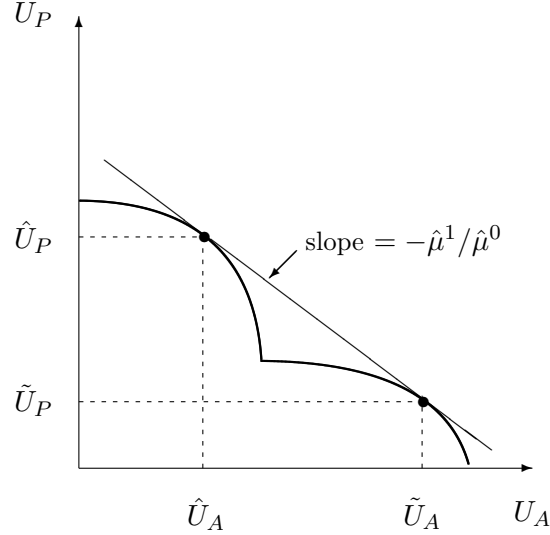
Figure 1: The case of a non-concave utility possibility frontier

Given that concavity does not guarantee that the MM approach works correctly one has to expect that the method fails also in the nonconcave case. In the appendix we provide a worked out example which shows this formally. In order to get an intuition for why in the nonconcave case the same problems arise as in our linear example consider Figure 4.2. Suppose a nonconcave optimization problem gives rise (at some state) to a (nonconvex) utility possibility set as depicted in this figure. Assume that the (unique) true solution of the problem is to choose in each period $\hat{U}_A$ as continuation utility for the agent. As one can see from the figure, the relative welfare weight corresponding to this point on the utility possibility frontier, $\hat{\mu}^1/\hat{\mu}^0$ supports also the choice of $\tilde{U}_A$. Hence, again there are multiple continuation utilities corresponding to one and the same relative welfare weight. Not being able to discriminate between the two points the the recursive saddle point method will admit for wrong solutions.

This discussion suggests that the MM method can work only if the sets of continuation utilities of the optimization problem are strictly convex, for only in that case any pair of welfare weights supports exactly one pair of continuation payoffs. Unfortunately, strict convexity of the utility possibility frontiers can only be guaranteed by imposing strict concavity assumptions on the problem under consideration. Obviously such a requirement is very costly, in that it reduces substantially the class of economic problems to which the approach might be applied. In fact MM themselves point out that problems which include incentive constraints often have a nonconcave structure.

These observations are most relevant for the literature mentioned in the introduction. As we have already remarked there, in most of those articles (Attanasio and

Rios-Rull, 2001, are an exception) the MM approach is applied to non-concave optimization problems. Our results show though that the application of the method to such problems is not very advisable.

It should be clear that the main argument of the present paper does not rely on the specificity of the example we are analyzing. Nevertheless, we would like to mention that the linear case is a very relevant one for two main reasons. First, economists very often use randomized allocations in order to convexify the utility possibility set of non-concave problems. Therefore a utility possibility frontier containing linear pieces is all but a degenerate situation.[15] In fact, the strategy to use randomized allocations has already been applied in papers using the MM technique. For instance, we can mention here the work of Ezra Friedman (1998). Second, in dynamic optimization problems the time dimension provides an 'implicit' instrument for the convexification of the utility possibility set. Hence, non concave problems may exhibit linear pieces in the utility possibility frontier even without explicit randomizations.

5   Conclusion

In this paper we have applied the recursive saddle point method developed by Marcet and Marimon (1999) to a simple concave dynamic optimization problem. Our results show that for problems which are not strictly concave, the conditions delivered by the MM method are not sufficient for optimality. It is interesting to notice that in the case of our example it turns out that the MM conditions are necessary.[16]

Given our findings it is natural to ask whether the reliability of solutions obtained by the MM approach could be guaranteed by ex-post checks of optimality and feasibility. Unfortunately, in general that seems to be hardly a viable procedure. In our example, for instance, that would require to check an infinite countable number of constraints. Moreover, to the unique optimal value $W^*(0,1)$ there corresponds a continuum of possible returns for the principal.

---

[15]There is a long list of references here. One of the earlier papers which proposes the use of lotteries in dynamic contracting is perhaps Phelan and Townsend (1991). Among the most recent contributions that have been using appropriate randomizations to convexify the problem one can name Ligon, Thomas and Worral (2000), Phelan and Stacchetti (2001), Albuquerque and Hopenhayn (2002), Clementi and Hopenhayn (2002). Ligon et. al. (2000) and Albuquerque and Hopenhayn (2002) are of particular interest, as their models fall into the MM-class, although they use the primal method.

[16]We do not know though how general this result is. We guess that the MM conditions are in general necessary for sequential saddle points. This would imply that they are also satisfied by every optimum whenever the problem under consideration is concave. On the other hand we can not draw such a conclusion for nonconcave problems since the conditions characterizing the set of sequential saddle points are only sufficient for a solution.

On the other hand, sometimes there might be specific situations where the *feasibility* check can be done - at least in an approximate way - in a finite number of steps. For example, consider the case where it is known that the optimal plan tends to a stationary set which is easy to identify and lies in the interior of the feasibility set.[17] Assume furthermore that the outcome of the MM procedure converges to the mentioned feasible stationary set. In this situation, one might expect that after a possibly large but finite number of periods feasibility is guaranteed, since the system is sufficiently 'close' to the feasible stationary set. As a consequence, one could restrict the check of feasibility to the initial periods of the transition.

Unfortunately, finding a general approach for the ex-post check of *optimality* seems even harder than the check for feasibility. The main reason of concern is that - as our example emphasizes - the MM procedure might lead to a continuum of possible candidates for the optimum.

---

[17]This seems to be the case of the 'unrestricted' steady state in the imperfect enforceability model of Marcet and Marimon (1992).

6 References

ABRAHAM, A. AND N. PAVONI (2004): "Efficient Allocations with Moral Hazard and Hidden Borrowing and Lending"; *mimeo*, UCL and Duke University.

ABREU, D., AND E. STACCHETTI (1990): "Toward a Theory of Discounted Repeated Games with Imperfect Monitoring"; *Econometrica*, Vol. 58(5), 1041-1064.

ALBUQUERQUE, R. AND H. HOPENHAYN (2002): "Optimal Lending Contracts and Firm Dynamics"; *mimeo*.

ATTANASIO, O. AND J. V. RIOS-RULL (2000):"Consumption Smoothing in Island Economies: Can Public Insurance Reduce Welfare?" *European Economic Review*, Vol. 44, 1225-1258.

AYAGARI, R., A. MARCET, TH. J. SARGENT AND J. SEPPÄLÄ (2002): "Optimal Taxation without State-Contingent Debt"; *Journal of Political Economy*, Vol. 110(6), 1220-1254.

CHANG, R. (1998) "Credible Monetary Policy in an Infinite Horizon Model: Recursive Approaches," *Journal of Economic Theory,"* 81: 431-461.

CLEMENTI, G. L. AND H. HOPENHAYN (2002): "A Theory of Financing Constraints and Firm Dynamics"; *Rochester Center For Economic Research*, WP # 492.

COOLEY, TH., R. MARIMON AND V. QUADRINI (2001): "Aggregate Consequences of Limited Contract Enforceability"; *mimeo*.

FRIEDMAN, E. (1998): "Risk Sharing and the Dynamics of Inequality"; *mimeo*.

KEHOE, P. AND F. PERRI (2002): "International Business Cycles with Endogenous Incomplete Markets"; *Econometrica*, Vol. 70(3), 907-928.

KHAN, A., R. G. KING AND A. L. WOLMAN (2002): "Optimal Monetary Policy"; *Federal Reserve Bank of Philadelphia Working Paper* No. 01-5.

KLEIN, P. AND J. V. RIOS-RULL (2002): "Time-Consistent Optimal Fiscal Policy"; Forthcoming *International Economic Review.*

LIGON, E., J. P. THOMAS, AND T. WORALL(2000): "Mutual Insurance, Individual Savings and Limited Commitment" *Review of Economic Dynamics*, Vol. 3(2), 216-246.

LUENBERGER, D. G. (1969): "Optimization by Vector Space Methods"; New York: Wiley.

LUCAS, R. E., AND N. L. STOCKEY (1983): "Optimal Fiscal and Monetary Policy in an Economy without Capital", *Journal of Monetary Economics.* Vol. 12, 55-93.

LUSTIG, H. AND S. VAN NIEUWERBURGH (2002): "Housing Collateral, Consumption Insurance and Risk Premia"; *mimeo.*

MARCET, A. AND R. MARIMON (1992): "Communication, Commitment and Growth"; *Journal of Economic Theory*, Vol. 58, 853-881.

MARCET, A. AND R. MARIMON (1999): "Recursive Contracts"; *Universitat Pompeu Fabra WP # 337.*

PHELAN, CH. AND E. STACCHETTI (2001): "Sequential Equilibria in a Ramsey Tax Model" *Econometrica* , Vol. 69(6), 1491-1518.

PHELAN, CH. AND R. M. TOWNSEND (1991): "Computing Multi-Period, Information-Constrained Optima" *Review of Economic Studies* , Vol. 58, 853-881.

SEPPÄLÄ, J. (1999): "Asset Prices and Business Cycles under Limited Commitment"; *mimeo.*

SIU, H. (2002): "Optimal fiscal and monetary policy with sticky prices"; *mimeo.*

STOCKEY, N. L., R. E. LUCAS, WITH E. C. PRESCOTT (1989): "Recursive Methods in Economics Dynamics"; Cambridge: Harvard University Press.

7 Appendix

In this appendix we show that the problem which arises in our linear example might be present also in the nonconcave case. In particular, we will consider again problem (1a)-(1c) and assume that $u(a) = a^2$, $\bar{a} = 1$ and $b = \beta^n$, where n is some natural number.

Before we proceed to solve this problem with the MM method, we would like to point out, that its unique 'true' solution is given by a stream of transfers which are equal to zero for the first $2n$ periods and equal to one afterwards. It is also important to notice that this optimal stream together with the sequence of Lagrange multipliers $(1, 0, 0, ...)$ constitutes a saddle point of the problem's associated Lagrangean (hence the failure of the MM method does not derive from non-existence of a sequential saddle point).

The recursive saddle point functional equation corresponding to the modified problem is given by

$$W(\mu^0, \mu^1) = \sup_{a \in [0, \bar{a}]} \inf_{\substack{\gamma^0, \gamma^1 \geq 0, \\ \mu^{0'}, \mu^{1'} \geq 0}} \mu^0 (y - a) + \gamma^0 (y - a - R) + \mu^1 a^2 + \gamma^1 (a^2 - \frac{b^2}{1 - \beta}) + \beta W(\mu^{0'}, \mu^{1'})$$

$$s.t. \quad \mu^{0'} = \mu^0 + \gamma^0$$

$$\mu^{1'} = \mu^1 + \gamma^1 \tag{11}$$

We will show in the following that the solution to this problem is given by

$$W^*(\mu^0, \mu^1) = \begin{cases} \mu^0 \frac{y - b^2}{1 - \beta} + \mu^1 \frac{b^2}{1 - \beta} & \text{if } \mu^0 \geq \mu^1 \\ \mu^0 \frac{y - 1}{1 - \beta} + \mu^1 \frac{1}{1 - \beta} & \text{if } \mu^0 < \mu^1 \end{cases}.$$

Furthermore, we will prove that the corresponding policy correspondences are

$$\phi_a(\mu^0, \mu^1) = \begin{cases} 1 & \text{if } \mu^0 < \mu^1 \\ \{0, 1\} & \text{if } \mu^0 = \mu^1 \\ 0 & \text{if } \mu^0 > \mu^1 \end{cases}$$

and

$$\phi_\gamma(\mu^0, \mu^1) = \begin{cases} (0, 0) & \text{if } \mu^0 < \mu^1 \\ (0, \mu^0 - \mu^1) & \text{if } \mu^0 \geq \mu^1. \end{cases}$$

Before we continue with the proof, notice that when starting with the initial values $\mu^0 = 1$ and $\mu^1 = 0$ the set of solutions generated with the above policy correspondences is simply the set of all sequences composed of zeros and ones which start with 0. In particular, this set contains also the zero sequence which of course fails to satisfy feasibility for any $1 > b > 0$. Hence, just as in the linear case the MM method allows for non-feasible transfer streams as solutions.

The following proof essentially follows the same steps as the proof for the linear case contained in the text. That is, we will simply show that the above policy correspondences solve the saddle point problem given W^*. Showing that W^* indeed then solves the functional equation is a simple algebraic exercise and is therefore omitted. We start by substituting the constraints in our saddle point problem which yields

$$\sup_{a\in[0,1]} \inf_{\gamma^0,\gamma^1\geq 0} \mu^0(y-a) + \gamma^0(y-a) + \mu^1 a^2 + \gamma^1(a^2 - \frac{b^2}{1-\beta}) + \beta W^*(\mu^0+\gamma^0, \mu^1+\gamma^1). \quad (12)$$

One can immediately see that the above expression is strictly increasing in γ_0, independently of a and the welfare weights μ^0 and μ^1. Hence, γ_0 must be equal to zero.

Next consider the choice of a. The only terms of (12) which depend on a are

$$-\mu^0 a + (\mu^1+\gamma^1)a^2.$$

This expression is convex in a. Hence, the only two candidates for a maximizer are the extreme points of the set of technically feasible transfers (i.e. $a=0$ and $a=1$). Which of the two candidates is optimal depends on whether μ^0 is larger or smaller than $\mu^1+\gamma^1$. While in the former case we get $a=0$ the latter case implies $a=1$ and of course both candidates are optimal if $\mu^0 = \mu^1+\gamma^1$.

Suppose now that $\mu^0 < \mu^1$. As γ^1 can never be negative we know that in this case we must always have $a=1$ (since the condition $\mu^0 < \mu^1 + \gamma_1$ is always satisfied). Therefore the slope of the objective in γ^1 is given by

$$1 - \frac{b^2}{1-\beta} + \frac{\beta}{1-\beta} = \frac{1-b^2}{1-\beta} > 0,$$

which implies that the unique minimizing value for γ^1 is zero.

Next, consider the case $\mu^0 > \mu^1$. We have to show that there is only one saddle point, namely $a=0$ and $\gamma^1 = \mu^0 - \mu^1$. In order to do so we will first prove that $\mu^0 - \mu^1$ is the unique minimizer of the objective in γ^1 for $a=0$. Since we have already seen that $a=0$ maximiezes (12) for $\gamma^1 = \mu^0 - \mu^1$ it then only remains to show that there is no saddle point with $a=1$.

Notice first that the slope of (12) in γ^1 is given by[18]

$$a^2 - \frac{b^2}{1-\beta} + \beta \begin{cases} \dfrac{b^2}{1-\beta} & \text{if } \gamma^1 \leq \mu^0 - \mu^1 \\ \dfrac{1}{1-\beta} & \text{if } \gamma^1 \geq \mu^0 - \mu^1. \end{cases} \quad (13)$$

Plugging in $a=0$ yields

$$-\frac{b^2}{1-\beta} + \beta\frac{b^2}{1-\beta} = -b^2 < 0$$

[18]Again, the slopes at the point $\gamma^1 = \mu^0 - \mu^1$ are to interpreted as lefthand and righthand derivatives respectively.

for the range $[0, \mu^0 - \mu^1]$ and

$$-\frac{b^2}{1-\beta} + \beta\frac{1}{1-\beta} = \frac{\beta - b^2}{1-\beta} > 0$$

if $\gamma^1 \geq \mu^0 - \mu^1$. Hence, if $a = 0$ (12) is indeed decreasing in γ^1 up to $\mu^0 - \mu^1$ and increasing afterwards (remember that $b = \beta^n$).

If instead $a = 1$ the slopes over the two ranges are

$$1 - \frac{b^2}{1-\beta} + \beta\frac{b^2}{1-\beta} = 1 - b^2$$

(over $[0, \mu^0 - \mu^1]$) and

$$1 - \frac{b^2}{1-\beta} + \beta\frac{1}{1-\beta} = \frac{1 - b^2}{1-\beta} \qquad (14)$$

(for $\gamma^1 \geq \mu^0 - \mu^1$). Both expressions are positive (since $b < 1$) and hence γ^1 would have to be chosen equal to zero. We have already seen though that $\gamma^1 < \mu^0 - \mu^1$ is only compatible with $a = 0$.

Finally, consider the case $\mu^0 = \mu^1$. It is easily verified that $\mu^0 - \mu^1 = 0$ is still the unique minimizer of (12) if $a = 0$ (and so $a = 0$, $\gamma^1 = \mu^0 - \mu^1 = 0$ is a saddle). But since in this situation γ^1 can never be chosen smaller then $\mu^0 - \mu^1$ (14) implies that also $a = 1$ and $\gamma^1 = 0$ is a saddle point, which concludes the prove.