The Ramsey Growth Model: Extensions and Algorithmic Solution

Karel Sladký

Department of Econometrics
Institute of Information Theory and Automation
Academy of Sciences of the Czech Republic
Pod Vodárenskou věží 4, 182 08 Praha 8
sladky@utia.cas.cz

Abstract

In this note, we consider in discrete-time finite state approximations of an extended Ramsey type model under stochastic uncertainty. Recalling standard procedure of stochastic dynamic programming we present explicit formulas for finding maximum global utility of the consumers (i.e. sum of total discounted instantaneous utilities) in the approximated model along with error bounds of the approximations.

Keywords

Economic dynamics, stochastic Ramsey type model, Markov decision chains, discounted optimality, discretization, error bounds

JEL Classification: C61, E21, E22

1 Classical Ramsey Model in Discrete-Time Setting

1.1 Formulation and Notations

The heart of the seminal paper of F. Ramsey [10] on mathematical theory of saving is an economy producing output from labour and capital and the task is to decide how to divide production between consumption and capital accumulation to maximize the global utility of the consumers. In contrast to the Ramsey's formulation considering the problem in continuous-time setting, in this note we work with a discrete-time version of the growth model.

Since in [10] the problem was considered in continuous-time setting, Ramsey suggested some variational methods for finding an optimal policy how to divide the production between consumption and capital accumulation.

Considering the Ramsey problem in discrete-time setting, the respective mathematical model can be formulated as follows:

We consider at discrete-time points $t = 0, 1, \dots$ an economy in which at each time t there are L_t identical consumers with consumption c_t per consumer. The number of consumers grow very slowly in time, i.e. $L_t = L_0(1+n)^t$ for $t = 0, 1, \dots$ with $\alpha := (1+n) \approx 1$. The economy produces at times $t = 0, 1, \dots$ gross output $\tilde{Y}_t$ using only two inputs: capital K_t and labour $L_t = L_0(1+n)^t$. A production function $F(K_t, L_t)$ relates input to output, i.e.

$$\tilde{Y}_t = F(K_t, L_t) \quad \text{with} \quad K_0 > 0,\ L_0 > 0 \ \text{ given.} \tag{1}$$

In each period output must be split between consumption $c_t L_t$ and gross investment I_t, i.e.

$$c_t L_t + I_t \le \tilde{Y}_t = F(K_t, L_t), \tag{2}$$

investment I_t is used in whole (along with the depreciated capital K_t) for the capital K_{t+1}. In addition, capital is assumed to depreciate at a constant rate $\delta \in (0,1)$, so capital related to gross investment at time point $t+1$ is equal to

$$K_{t+1} = (1-\delta)K_t + I_t. \tag{3}$$

Preferences over consumption of a single consumer (resp. the considered L_t consumers) are taken to be of the form

$$\sum_{t=0}^{T} \beta^t u(c_t) \quad \left(\text{resp. } L_0 \sum_{t=0}^{T} (\alpha\beta)^t u(c_t)\right) \quad \text{for a finite or infinite time horizon } T, \tag{4}$$

where $u(\cdot)$ is instantaneous utility function and $\beta < 1$ is a given discount factor. Observe that if $\alpha\beta < 1$ then $L_0 \sum_{t=0}^{\infty} (\alpha\beta)^t u(c_t) < \infty$.

The problem is to find the rule how to split production between consumption and capital accumulation that maximizes global utility of the consumers for a finite or infinite time horizon T.

Denoting by $k_t := K_t/L_t$ the capital per consumer at time t, and similarly by $y_t := Y_t/L_t = F(k_t, 1)$ the output per consumer at time t (note that $F(\cdot,\cdot)$ is assumed to be homogeneous of degree one, i.e. $F(\theta K, \theta L) = \theta F(K, L)$ for any $\theta \in \mathbb{R}$), from (2), (3) we get

$$c_t + k_{t+1} - (1-\delta)k_t \le y_t = F(k_t, 1), \tag{5}$$

and if we define the function $f(\cdot)$ by $f(k) := F(k,1) + (1-\delta)k$ then (5) can be written as

$$c_t + k_{t+1} \le y_t = f(k_t) \tag{6}$$

where y_t is the total output at time t.

In the above formulation we assume that the production function $f(k)$ and the consumption function $u(c)$ fulfil some standard assumptions on production and consumption functions, in particular, that:

AS 1. The function $u(c) : \mathbb{R}^+ \to \mathbb{R}^+$ is twice continuously differentiable and satisfies $u(0) = 0$. Moreover, $u(c)$ is strictly increasing and concave (i.e., its derivatives satisfy $u'(\cdot) > 0$ and $u''(\cdot) < 0$) with $u'(0) = +\infty$ (so-called Inada Condition).

AS 2. The function $f(k) : \mathbb{R}^+ \to \mathbb{R}^+$ is twice continuously differentiable and satisfies $f(0) = 0$. Moreover, $f(k)$ is strictly increasing and concave (i.e., its derivatives satisfy $f'(\cdot) > 0$ and $f''(\cdot) < 0$) with $f'(0) = M < +\infty$, $\lim_{k\to\infty} f'(k) < 1$.

This note presents a slight extension of the author's paper [13] in which the main ideas concerning discretization and further extension of the stochastic version of the Ramsey model were reported. However, in the present paper computational aspects of the discretized stochastic models are investigated in detail; in particular, algorithmic procedures along with error bound and sensitivity analysis are elaborated.

1.2 Finding Optimal Policies by Dynamic Programming

Finding a sequence $(\boldsymbol{k}, \boldsymbol{c})^T = \{k_0, c_0, k_1, c_1, \dots, k_T, c_T\}$ with a given $k_0 > 0$ maximizing (4) under condition (6) can be formulated as:

Find

$$\hat{U}^{\beta}_{k_0}(T) := \max_{(\boldsymbol{k},\boldsymbol{c})^T} \sum_{t=0}^{T} \beta^t u(c_t) \quad \text{for a finite or infinite time horizon } T, \tag{7}$$

under the constraints (for $t = 0, 1, \dots, T$)

$$c_t + k_{t+1} \le f(k_t) \tag{8}$$

$$c_t \ge 0, \quad k_t \ge 0, \quad \text{with } k_0 > 0 \text{ given.} \tag{9}$$

Note that since $u(\cdot)$, $f(\cdot)$ are increasing (cf. assumptions AS 1 and AS 2) it is possible to replace the constraints (8), (9) by

$$c_t + k_{t+1} = f(k_t), \quad \text{with} \quad f(k_t) = y_t \tag{10}$$

$$c_t \ge 0, \quad k_t \ge 0, \quad k_0 > 0 \text{ given, and if } T < +\infty \text{ also } k_{T+1} = 0, \tag{11}$$

and hence also (7) can be written as

$$\hat{U}^{\beta}_{k_0}(T) = \max_{\boldsymbol{k}^T} \sum_{t=0}^{T} \beta^t u(f(k_t) - k_{t+1}) \quad \text{for a finite or infinite time horizon } T, \tag{12}$$

where $\boldsymbol{k}^T = \{k_0, k_1, \dots, k_T\}$ and $U^{\beta,\boldsymbol{k}^T}_{k_0}(T) = \sum_{t=0}^{T} \beta^t u(f(k_t) - k_{t+1})$.

Observe that in virtue of assumption AS 2 and (8), (10) it holds:

Fact 1. i) If $f'(0) \le 1$ (and hence $f'(k) < 1$ for all $k > 0$), then by (10) every sequence $\{k_0, k_1, \dots, k_t, \dots\}$ must be decreasing and $\lim_{t\to\infty} k_t = 0$.

ii) If $f'(0) > 1$ (and hence, since $\lim_{k\to\infty} f'(k) < 1$, there exists some k' such that $f'(k) < 1$ for all $k > k'$), then there exists some $k^* > 0$ such that $f(k^*) = k^*$ and some $k_m \in (0, k^*)$ such that $f(k_m) - k_m = \max_k[f(k) - k]$.
Supposing that $k_0 > k^*$ then elements of the sequence $\{k_0, k_1, \ldots, k_t, \ldots\}$ must be decreasing for all $k_t > k^*$. Furthermore, if for some $t = t_\ell$ it holds $k_{t_\ell} < k^*$ then $k_t < k^*$ for all $t \geq t_\ell$, but $\{k_t, t \geq t_\ell\}$ need not be monotonous. However, in any case $k_t \leq k_{\max} = \max(k_0, k^*)$ and $f(k_t) \leq f(k_{\max}) =: y_{\max}$ for all $t = 0, 1, \ldots$.

iii) In case that $k_0' > k_0 > 0$ then $\hat{U}^\beta_{k_0}(T) > \hat{U}^\beta_{k_0'}(T)$. This can be easily verified since if we start with capital $k_0' > k_0$, selecting consumption at time 0 such that $c_0' + k_1 = f(k_0') > c_0 + k_1 = f(k_0)$ (recall that $f(k_0') > f(k_0)$ and $u(\cdot)$ is increasing) and following for every $t > 0$ decisions given by $\boldsymbol{k}^T \equiv (k_0, k_1, \ldots, k_T)$ (the sequence of capital stocks yielding $\hat{U}^\beta_{k_0}(T)$ in (12)), then $u(f(k_0') - k_1) > u(f(k_0) - k_1)$ and $u(f(k_t') - k_{t+1}')$ with $k_t' \equiv k_t$ for all $t \geq 1$).

On employing separability occurring in (12), for finite T and a given $k_0 > 0$ we get

$$\begin{aligned}
\hat{U}^\beta_{k_0}(T) &= \max_{k_1}\left[u(f(k_0) - k_1) + \beta \hat{U}^\beta_{k_1}(T-1)\right],\\
\hat{U}^\beta_{k_1}(T-1) &= \max_{k_2}\left[u(f(k_1) - k_2) + \beta \hat{U}^\beta_{k_2}(T-2)\right],\\
&\vdots\\
\hat{U}^\beta_{k_{T-2}}(2) &= \max_{k_{T-1}}\left[u(f(k_{T-2} - k_{T-1}) + \beta \hat{U}^\beta_{k_{T-1}}(1)\right],\\
\hat{U}^\beta_{k_{T-1}}(1) &= \max_{k_T}\left[u(f(k_T)\right],
\end{aligned}$$

and hence (using the celebrated Bellman's "principle of optimality", see, e.g. [1])

$$\begin{aligned}
\hat{U}^\beta_{k_0}(T) = {} & \max_{k_1}[u(f(k_1) - k_0) + \beta \max_{k_2}[u(f(k_2) - k_1)\\
& + \beta \max_{k_3}[u(f(k_3) - k_2) + \beta \max_{k_4}[u(f(k_4) - k_3)\\
& + \ldots + \beta \max_{k_{T-1}}[u(f(k_{T-2}), k_{T-1}) + \beta \max_{k_T} u(k_T)] \ldots]]]].
\end{aligned} \tag{13}$$

Now let us introduce the opposite time orientation, i.e. if T is fixed then for $n = 0, 1, \ldots, T$, let $c^n = c_{T-n}$, $k^n = k_{T-n}$. Then (13) can be rewritten as:

$$\begin{aligned}
\hat{U}^\beta_{k^T}(T) = {} & \max_{k^{T-1}}[u(f(k^{T-1}) - k^T) + \beta \max_{k^{T-2}}[u(f(k^{T-2}) - k^{T-1})\\
& + \beta \max_{k^{T-3}}[u(f(k^{T-3}) - k^{T-2}) + \beta \max_{k^{T-4}}[u(f(k^{T-4}) - k^{T-3})\\
& + \ldots + \beta \max_{k^1}[u(k^1, 1) + \beta \max_{k^0} u(k^0)] \ldots]]]]
\end{aligned} \tag{14}$$

or

$$\hat{U}^\beta_{k^n}(n) = \max_{k^{n-1}}[u(f(k^{n-1}) - k^n) + \beta \hat{U}^\beta_{k^{n-1}}(n-1)] \qquad \text{for } n = 1, 2, \ldots, T, \tag{15}$$

where $\hat{U}^\beta_{k^0}(0) = \max_{k^0} u(k^0)$ for the selected value $k^0 > 0$.

2 Economic Growth Under Uncertainty: Probabilistic Approach

2.1 Uncertainty Modelled by Markov Processes

Up to now we have assumed that for a given k_t the total output $y_t = f(k_t)$ is determined by (10). To include random shocks or imprecisions into the model, we shall assume that for a given value of k_t we obtain the output y_t only with known probability $p(k_t) \equiv p(k_t; 0) < 1$, hence with probability $\bar{p}(k_t) = 1 - p(k_t)$ the total output will be different from y_t and can attain maximal and minimal possible values $f_{\max}(k_t)$ and $f_{\min}(k_t)$ respectively (of course, we assume that assumptions AS 2 also hold for $f_{\max}(\cdot)$ and $f_{\min}(\cdot)$). Since (cf. assumption AS 1) instantaneous utility function $u(\cdot)$ is increasing, on *replacing the production function $f(k_t)$ by $f_{\max}(k_t)$ and $f_{\min}(k_t)$ we obtain upper or lower bounds on the total output at time t and also, for fixed values of k_t, also the upper and lower bounds on the maximal global utility of the consumers respectively.* This approach is relatively simple, but ignores a lot of information and yields only a very rough bounds on optimal values.

Obviously, significantly better results can be obtained if we replace the rough estimates of y_t generated by means of $f_{\max}(k_t)$ and $f_{\min}(k_t)$ by a more detailed information on the (random) output y_t generated by the capital k_t.

Recall that by Fact 1ii the values of k_t, $y_t = f(k_t)$ (and hence also c_t) are bounded by $k_{\max}$, $y_{\max}$ respectively with $f_{\max} := y_{\max}$.

To this end we shall assume that in (6), (10)

$$y_t = Z_t f(k_t), \text{ where } Z = \{Z_t, t = 0, 1, \dots\} \text{ is a random process.} \tag{16}$$

In the literature (cf. [5, 8] or the monograph [14]) it is usually assumed that Z is a Markov process (in general with state space $\mathbb{R}$) or an autoregressive process. Moreover, we assume that *the decision maker can observe the current values of the total output y_t and then select the value of k_{t+1}*. Such an extension well corresponds to the models introduced and studied in [14] and also in [5, 8]. Unfortunately, assuming that Z is a Markov process with compact state space $\mathbb{R}$ then a rigorous treatment of the model given by (15) requires a very sophisticated mathematics (see [3] or [14]) and is not suitable for numerical computation. To make the model computationally tractable we shall approximate our system governed by (10), (11) (with $f(\cdot) = f_{\max}(\cdot)$) by a discretized model with finite state space.

2.2 Discretized Markov Model

In what follows, we shall assume that the values of c_t, k_t, and y_t take on only discrete values. In particular, we assume that for sufficiently small $\Delta > 0$ there exists nonnegative integers $\bar{c}_t$, $\bar{k}_t$, and $\bar{y}_t$ such that for every $t = 0, 1, \dots$ it holds:
$\bar{c}_t \Delta = c_t$, $\bar{k}_t \Delta = k_t$, and $\bar{y}_t \Delta = y_t$ with $\bar{k}_t \le K := k_{\max}/\Delta$ and similarly $\bar{y}_t \le Y := y_{\max}/\Delta$. Let elements of $\bar{k}_t$ be labelled by integers from $\mathcal{I}_K = \{0, 1, \dots, K\}$ and elements of $\bar{y}_t$ by integers from $\mathcal{I}_Y = \{0, 1, \dots, Y\}$. Hence for the total output y_t generated by the "randomized" production function we get for $\ell = 0, 1, 2, \dots, L$

$$\bar{y}_t = f(\bar{k}_t) - \ell\Delta \quad \text{with known probability } p(\bar{k}_t; \ell); \text{ obviously, } \sum_{\ell=0}^{L} p(\bar{k}_t; \ell) = 1, \tag{17}$$

and let $\boldsymbol{p}(\bar{k}_t) = [p(\bar{k}_t; 0), p(\bar{k}_t; 1), \ldots, p(\bar{k}_t; L)]$.
We shall assume that $\boldsymbol{p}(\bar{k}_t)$ is "close" to $\boldsymbol{p}(\bar{k}_{t+1})$ for every $\bar{k}_t$, i.e. assume existence of some $\tilde{\Delta} > 0$ such that $|p(\bar{k}_{t+1}; \ell) - p(\bar{k}_t; \ell)| < \tilde{\Delta}$ for every $\ell = 1, \ldots, L$ and $\bar{k}_t = 1, \ldots, K$.

If the (random) total output at time t $\bar{y}_t = \bar{y}$ then the decision maker have option to invest for the next time point the capital $k_{t+1} = \bar{k}_{t+1}\Delta$ where $\bar{k}_{t+1} = \bar{g}_t, \ldots, \bar{y}_t$ (with given $\bar{g}_t = 0, 1, \ldots, f_{\max}(\bar{k}_t)$), and hence $\beta^t u((\bar{y} - \bar{k}_{t+1})\Delta)$ is the instantaneous utility accrued at time t to the global utility. In accordance with decision d taken at time t if the output $\bar{y}_t = \bar{y}$, at the next time $t+1$ capital $\bar{k}_{t+1}$ will be employed, see also the following diagram

$$\bar{k}_t \xrightarrow{p(\bar{k}_t, \bar{y}_t)} \bar{y}_t \xrightarrow{d} \bar{k}_{t+1}.$$

Using the above discretization and taking decisions with respect to the current states, the development of the economy over time can be well described by a (structured) Markov reward chain $X = \{X_\tau,\ \tau = 0, 1, \ldots\}$ with finite state space $\mathcal{I} = \mathcal{I}_K \cup \mathcal{I}_Y$ (with $\mathcal{I}_K \cap \mathcal{I}_Y = \emptyset$), transition probabilities $p(\bar{k}_t; \bar{y}_t) = p_{ij}$, for $i = \bar{k}_t \in \mathcal{I}_K$, $j = \bar{y}_t \in \mathcal{I}_Y$, and a "non-random" transition from state $j = \bar{y}_t \in \mathcal{I}_Y$ to state $\ell = \bar{k}_t \in \mathcal{I}_K$ *associated with one-stage reward* $r_{j\ell} = u((\bar{y}_t - \bar{k}_{t+1})\Delta)$. Observe that actually "two transitions" of the Markov chain $X = \{X_\tau, \tau = 0, 1, \ldots\}$ occur within one-time period of the considered economy model and the *one-stage reward is accrued only in even transitions.* Hence the global utility (i.e. the total discounted reward of the Markov chain X) $U^{\beta}_{\bar{k}_0}(T) = \mathsf{E}\{\sum_{t=1}^{T} \beta^t r_{X_{2t-1}, X_{2t}} | X_0 = \bar{k}_0\}$ (the symbol E is reserved for expectation).

3 Some Generalization of the Discretized Model

3.1 Extension of the Stochastic Growth Model

Up to now we have assumed that the probability vector $\boldsymbol{p}(\bar{k}_t)$ cannot be influenced by the decision maker. Now we extend the model in such a way that $\boldsymbol{p}(\bar{k}_t)$ will be replaced by a family of vectors $\boldsymbol{p}(\bar{k}_t, d(\bar{k}_t))$ for $d(\bar{k}_t) = 1, 2, \ldots, D$ depending on the decision taken in state $\bar{k}_t$. Moreover, some *cost,* denoted $c(d(\bar{k}_t))$, *will be accrued to this decision.*

Moreover, we shall assume that the decision d, taken if at time t the output $\bar{y}_t = \bar{y}$, assign the desired values of capital only with some probability, i.e. there is a set of feasible decisions $d(\bar{y}_t) = 1, 2, \ldots, D$ each of them assigns the value of capital $\bar{k}_{t+1}$ with known probability vector $\boldsymbol{p}(\bar{y}_t, d(\bar{y}_t)) = [p(\bar{k}_t; 1, d(\bar{y}_t)), p(\bar{k}_t; 2, d(\bar{y}_t)), \ldots, p(\bar{k}_t; \bar{y}_t, d(\bar{y}_t))]$.

We shall assume that $\boldsymbol{p}(\bar{k}_t, d)$ and $\boldsymbol{p}(\bar{y}_t, d)$ is "close" to $\boldsymbol{p}(\bar{k}_{t+1}, d)$ and to $\boldsymbol{p}(\bar{y}_{t+1}, d)$ for every $\bar{k}_t$ and $\bar{y}_t$ respectively, i.e. we assume existence of some $\tilde{\Delta} > 0$ such that $|p(\bar{k}_{t+1}; \ell, d) - p(\bar{k}_t; \ell, d)| < \tilde{\Delta}$ for every $\ell = 1, \ldots, L$ and $\bar{k}_t = 1, \ldots, K$ and $|p(\bar{y}_{t+1}; \ell, d) - p(\bar{y}_t; \ell, d)| < \tilde{\Delta}$ for every $\ell = 1, \ldots, D$ and $\bar{y}_t = 1, \ldots, K$. So the development over time is given by the following diagram

$$\bar{k}_t \xrightarrow[]{\substack{c(d(\bar{k}_t)) \\ p(\bar{k}_t, \bar{y}_t; d(\bar{k}_t))}} \bar{y}_t \xrightarrow[]{\substack{u((\bar{y}_t - \bar{k}_{t+1})\Delta) \\ p(\bar{y}_t, \bar{k}_{t+1}; d(\bar{y}_t))}} \bar{k}_{t+1}$$

In contrast to the previous model transition from state $\bar{y}_t \in \mathcal{I}_Y$ to state $\bar{k}_t \in \mathcal{I}_K$ is random and given by a known probability vector $\boldsymbol{p}(\bar{y}_t, d(\bar{y}_t))$ depending on the selected decision.

3.2 Formulation in Terms of Stochastic Dynamic Programming

The above model can be considered as a structured standard Markov decision chain with finite state space $\mathcal{I} = \mathcal{I}_1 \cup \mathcal{I}_2$ (with $\mathcal{I}_1 \cap \mathcal{I}_2 = \emptyset$), finite set $\mathcal{D}_i = \{0, 1, \ldots, d(i)\}$ of possible decisions (actions) in state $i \in \mathcal{I}$ and the following transition and reward structure:

$p_{ij}(a)$: transition probability from $i \to j$ $(i, j \in \mathcal{I})$ if action $a \in \mathcal{D}_i$ is selected,

r_{ij} : one-stage reward for a transition from $i \to j$, with
$r_{ij} = u((i-j)\Delta)$ if $i \in \mathcal{I}_2$ and $j \in \mathcal{I}_1$,
$r_{ij} = c(a)$ if $i \in \mathcal{I}_1$ and $j \in \mathcal{I}_2$, and action a is selected,

$r_i(a)$: expected value of the one-stage rewards incurred in state i if decision (or action) $a \in \mathcal{D}_i$ is selected in state i; in particular $r_i(a) = \sum_{j \in \mathcal{I}} p_{ij}(a) \cdot r_{ij}$.

A policy controlling the chain, say π, is a rule how to select actions in each state. Policy π is then fully identified by a sequence $\{d_\tau, \tau = 0, 1, \ldots\}$ of decision vectors (of dimension K and Y in odd and even steps respectively) whose ith element $d_\tau(i) \in \mathcal{D}_i$ identifies the action taken if $X_\tau = i$. If we restrict on stationary policies, i.e. the rules selecting actions only with respect to the current state of Markov chain X, then policy π is fully determined by $d_t \equiv d$. Observe that decision vector d then completely identifies the transition probability matrix $P(d)$ and the ith row of $P(d)$ has elements $p_{i1}(d(i)), \ldots, p_{iN}(d(i))$. Similarly, $r(d)$ is a (column) vector of one-stage expected rewards (i.e. i-th element of $r(d)$ is equal to $r_i(d(i))$).

Let the vector $U^{\beta,\pi}(\tau)$ denote expectation of the (random) global utility ξ_τ received in the τ next transitions of the considered Markov chain X if policy $\pi = (d_\tau)$ is followed, given the initial state $X_0 = i$, i.e., for the elements of $U^{\beta,\pi}(\tau)$ we have $U_i^{\beta,\pi}(\tau) = \mathsf{E}_i^\pi[\xi_\tau]$ where $\xi_\tau = \sum_{k=0}^{\tau-1} \beta^k r_{X_k, X_{k+1}}$ and E_i^π is the expectation if $X_0 = i$ and policy $\pi = (d_\tau)$ is followed. Then obviously

$$U_i^{\beta,\pi}(\tau+1) = r_i(d_\tau(i)) + \beta \sum_{j \in \mathcal{I}} p_{ij}(d_\tau(i)) \cdot U_j^{\beta,\pi}(\tau), \quad i \in \mathcal{I} \tag{18}$$

and for t (or τ) tending to infinity, i.e. when $\lim_{\tau \to \infty} U_i^{\beta,\pi}(\tau) = U_i^{\beta,\pi}$, (18) takes on the form

$$U_i^{\beta,\pi} = r_i(d(i)) + \beta \sum_{j \in \mathcal{I}} p_{ij}(d(i)) \cdot U_j^{\beta,\pi}, \quad i \in \mathcal{I}. \tag{19}$$

If $\hat{\pi}^T$ is (in general nonstationary) policy maximizing the values $U_i^{\beta,\pi}(T)$ for the fixed time horizon T then

$$U_i^{\hat{\pi}^\tau}(\tau) = \max_{d \in \mathcal{D}_i}[r_i(d_\tau(i)) + \beta \sum_{j \in \mathcal{I}} p_{ij}(d_\tau(i)) \cdot U_j^{\hat{\pi}^{\tau-1}}(\tau - 1)], \quad \text{for } \tau = T, T-1, \ldots, 1, 0. \tag{20}$$

Furthermore, for T tending to infinity, i.e. when $\lim_{T \to \infty} U_i^{\hat{\pi}^T}(T) = U_i^{\hat{\pi}}$, then (19), (20) read

$$U_i^{\hat{\pi}} = \max_{d \in \mathcal{D}_i}[r_i(d(i)) + \beta \sum_{j \in \mathcal{I}} p_{ij}(d(i)) \cdot U_j^{\hat{\pi}}], \quad i \in \mathcal{I}. \tag{21}$$

3.3 Computation of Optimal Policies

In case that the time horizon T is finite, it is necessary to calculate (backwards) the dynamic programming recursion according to (20). Considering the infinite time horizon (i.e. if $T \to \infty$), finding a solution of (21) is in some aspects much easier (optimal policy can be found in the class of stationary policies (i.e. policies selecting actions only with respect to the current state of Markov chain) and can be performed either by value iterations (successive approximations) or by policy iterations.

Algorithm 1 (Policy iterations).

Step 0. Select arbitrary policy, say $d^{(0)}$.
Step 1 – Policy evaluation. For stationary policy $d^{(n)}$ find $v = v(d^{(n)})$ as the solution of

$$v = r(d^{(n)}) + \beta P(d^{(n)}).$$

Step 2 – Policy improvement. For a given $v(d^{(n)})$ find policy $d^{(n+1)}$ such that

$$r(d^{(n+1)}) + \beta P(d^{(n+1)})v(d^{(n)}) = \max_{d \in D}[r(f) + \beta P(d)v(d^{(n)})].$$

If there exists $d^{(n+1)} = d^{(n)}$, then stop and policy $d^{(n)}$ is an optimal policy, else go to Step 1.

Algorithm 2 (Value iteration).

Select $v^{(0)} = 0$, choose some (sufficiently small) $\varepsilon > 0$, and iterate

$$v^{(n+1)} := \max_{d \in D}[r(d) + \beta P(d)v^{(n)}] \quad \text{for} \quad n = 0, 1, \ldots.$$

If $\|v^{(n+1)} - v^{(n)}\| < \varepsilon$ then stop.

Remark. Observe that $v^{(n)}$'s are identical with $U^{\beta,\pi}(n)$ if policy π is identified by the decision vectors generated by Algorithm 2.

Algorithm 3 (Value iteration (modified)).

Select $w^{(0)} = 0$, choose some (sufficiently small) $\varepsilon > 0$, set $w_N^{(n)} \equiv 0$ for $n = 0, 1, \ldots$, and iterate

$$w^{(n+1)} := \max_{d \in D}[r(d) + \beta P(d)w^{(n)} - (1 - \beta)v_N^{(n)}].$$

If $\|w^{(n+1)} - w^{(n)}\| < \varepsilon$ then stop.

3.4 Error Bounds

3.4.1 Rough Bounds

Observe that in Section 2.2 the difference of all data in the original and discretized model must be nongreater than the considered values $\Delta > 0$.

Since the difference between the original and discretized discounted rewards accrued at the stage n is is nongreater than $\beta^n \cdot \Delta$, taking into account only the considered discounting, the error caused by approximations in the finite horizon model (with n transitions), resp. in the infinite horizon model, is obviously bounded

$$\text{by} \quad \Delta \cdot \frac{1-\beta^n}{1-\beta} \qquad \text{resp.} \quad \text{by} \quad \Delta \cdot \frac{1}{1-\beta}.$$

Of course, the above bounds do not take into consideration transition probabilities as well as imprecisions in transition probabilities arising by discretization.

3.4.2 Finer Bounds

Taking into account also the errors arising by discretization in the transition probability matrix, for the original model, denoted (P, r), and the discretized model, denoted $(\bar{P}, \bar{r})$ we immediately get (for the sake of brevity we consider only stationary policies, and suppose that $r, \bar{r} \geq 0$; I denotes an identity matrix)

$$\begin{aligned} \bar{v} &= \sum_{k=0}^{\infty} (\beta \bar{P})^k \cdot \bar{r} = [I - \beta \bar{P}]^{-1} \cdot \bar{r} \\ \bar{v}^{(n)} &= \sum_{k=0}^{n-1} (\beta \bar{P})^k \cdot \bar{r} = \sum_{k=0}^{\infty} (\beta \bar{P})^k \cdot \bar{r} - (\beta \bar{P})^n \sum_{k=0}^{\infty} (\beta \bar{P})^k \cdot \bar{r}. \end{aligned}$$

However, since $r - \Delta \cdot e \leq \bar{r} \leq r + \Delta \cdot e$ (e denotes a unit vector)

$$\begin{aligned} \bar{v} &\leq [I - \beta(P + \Delta I)]^{-1} \cdot \bar{r} + \Delta \cdot [I - \beta(P + \Delta I)]^{-1} \cdot e \\ \bar{v} &\geq [I - \beta(P - \Delta I)]^{-1} \cdot \bar{r} - \Delta \cdot [I - \beta(P - \Delta I)]^{-1} \cdot e \end{aligned}$$

and the inverse of the matrix $[I - \beta \bar{P}]$ can be well approximated e.g. by

$$[I - \beta(P \pm \Delta I)]^{-1} \approx [I - \beta P]^{-1} \pm \Delta \cdot \frac{\beta}{1-\beta} \cdot [I - \beta P]^{-1}$$

or by

$$[I - \beta(P \pm \Delta I)]^{-1} \approx [I - \beta P]^{-1} \cdot \{I \pm \Delta\beta \cdot [I - \beta P]^{-1}\}.$$

(The above formulae follow on expanding $[I - \beta(P \pm \Delta I)]^{-1}$ and collecting the terms.)

Acknowledgement. This research was supported by the Czech Science Foundation under Grants 402/05/0115 and 402/06/0990.

References

[1] R. Bellman: Dynamic Programming. Princenton Univ. Press, Princenton, NJ. 1957.

[2] D. Blackwell: Discrete dynamic programming. Ann. Math. Statist. *33* (1962), 719–726.

[3] D. Blackwell: Discounted dynamic programming. Ann. Math. Statist. *36* (1965), 226–235.

[4] R.-A. Dana and C. Le Van: Dynamic Programming in Economics. Kluwer, Dordrecht 2003.

[5] B. Heer and A. Maußer: Dynamic General Equilibrium Modelling. Springer–Verlag, Berlin 2005.

[6] R. A. Howard: Dynamic Programming and Markov Processes. MIT Press, Cambridge, Mass. 1960.

[7] J. MacQueen: A modified dynamic programming method of successive approximations. Journal Math. Anal. Applic. *14* (1966), 38–43.

[8] M. Majamdar, T. Mitra and K. Nishinmura (eds.): Optimization and Chaos. Springer–Verlag, Berlin 2000.

[9] M. L. Puterman: Markov Decision Processes – Discrete Stochastic Dynamic Programming. Wiley, New York 1994.

[10] F. P. Ramsey: A mathematical theory of saving. Economic Journal *38* (1928), 543–559.

[11] S. M. Ross: Applied Probability Models with Optimization Applications. Holden–Day, San Francisco, Calif. 1970.

[12] K. Sladký: Sensitivity formulas for discrete- and continuous-time Markov chians. In: Prague Stochastics'98. Union of the Czech Mathematicians and Physicists, Prague 1998 pp. 517–521.

[13] K. Sladký: Approximations in stochastic growth models. In: Proceedings of the 24th Internat. Conference Mathematical Methods in Economics 2006, University of West Bohemia, Pilsen 2006, pp. 465–470.

[14] N.-L. Stokey and R. E. Lukas, Jr.: Recursive Methods in Economic Dynamics. Harvard Univ. Press, Cambridge, Mass. 1989.