Karel Sladký

Growth rates and average optimality in risk-sensitive Markov decision chains

Kybernetika, Vol. 44 (2008), No. 2, 205--226

Persistent URL: http://dml.cz/dmlcz/135844

KYBERNETIKA — VOLUME *44* (2008), NUMBER 2, PAGES 205 – 226

GROWTH RATES AND AVERAGE OPTIMALITY IN RISK–SENSITIVE MARKOV DECISION CHAINS

Karel Sladký

In this note we focus attention on characterizations of policies maximizing growth rate of expected utility, along with average of the associated certainty equivalent, in risk-sensitive Markov decision chains with finite state and action spaces. In contrast to the existing literature the problem is handled by methods of stochastic dynamic programming on condition that the transition probabilities are replaced by general nonnegative matrices. Using the block-triangular decomposition of a collection of nonnegative matrices we establish necessary and sufficient conditions guaranteeing independence of optimal values on starting state along with partition of the state space into subsets with constant optimal values. Finally, for models with growth rate independent of the starting state we show how the methods work if we minimize growth rate or average of the certainty equivalent.

Keywords: risk-sensitive Markov decision chains, average optimal policies, optimal growth rates

AMS Subject Classification: 90C40, 60J10, 93E20

1. INTRODUCTION AND NOTATION

In recent years there is a growing interest in so called risk-sensitive Markov decision processes when the outcome, say ξ, generated by a Markov reward process is evaluated using a utility function with constant risk sensitivity $\gamma \in \mathbb{R}$ (see e. g. [4, 5, 6, 7, 8, 9, 10]). Then the utility function $u^\gamma(\cdot)$ takes on the following form

$$u^\gamma(\xi) := \begin{cases} \operatorname{sign}(\gamma)\, \exp(\gamma\xi), & \text{if } \gamma \neq 0 \\ \xi & \text{for } \gamma = 0. \end{cases} \tag{1.1}$$

Obviously $u^\gamma(\cdot)$ is continuous, strictly increasing, and convex (resp. concave) for $\gamma > 0$, the risk seeking case (resp. $\gamma < 0$, the risk averse case). In case that $\gamma = 0$ the utility function is risk-neutral, i. e. a linear function and neither large nor small values of ξ are preferred.

The aim of this note is a complete characterization of policies maximizing growth rate of expected utility, along with average of the associated certainty equivalent, in risk-sensitive Markov decision chains with finite state and action spaces.

We consider a Markov decision chain $X = \{X_n,\, n = 0, 1, \ldots\}$ with finite state space $\mathcal{I} = \{1, 2, \ldots, N\}$ and finite set $\mathcal{A}_i = \{1, 2, \ldots, K_i\}$ of possible decisions (actions) in state $i \in \mathcal{I}$. Supposing that in state $i \in \mathcal{I}$ action $a \in \mathcal{A}_i$ is selected, then state j is reached in the next transition with a given probability $p_{ij}(a)$ and one-stage transition reward $r_{ij}(a)$ will be accrued to such transition.

A (Markovian) policy controlling the chain, $\pi = (f^0, f^1, \ldots)$, is identified by a sequence of decision vectors $\{f^n, n = 0, 1, \ldots\}$ where $f^n \in \mathcal{F} \equiv \mathcal{A}_1 \times \ldots \times \mathcal{A}_N$ for every $n = 0, 1, 2, \ldots$, and $f^n_i \in \mathcal{A}_i$ is the decision (or action) taken at the nth transition if the chain X is in state i. Policy which takes at all times the same decision rule, i. e. $\pi \sim (f)$, is called stationary; $\boldsymbol{P}(f)$ is transition probability matrix with elements $p_{ij}(f_i)$. Let $\xi^n_{X_0}(\pi) = \sum_{k=0}^{n-1} r_{X_k,X_{k+1}}(f^k_{X_k})$ be the stream of transition rewards received in the n next transitions of the considered Markov chain X if policy $\pi = (f^n)$ is followed and the process starts in state X_0. Similarly, let $\xi^{(m,n)}_{X_m}(\pi)$ be the total (random) reward obtained from the mth up to the nth transition (obviously, $\xi^n_{X_0}(\pi) = \xi^{(0,n)}_{X_0}(\pi) = r_{X_0,X_1}(f^0_{X_0}) + \xi^{(1,n)}_{X_1}(\pi)$).

In this article we assume that the stream of transition rewards $\xi^n_{X_0}$ (for the sake of brevity we often delete the argument π) is evaluated by an exponential utility function given by (1.1). In particular, for the (random) utility assigned to $\xi^n_{X_0}$, we have

$$u^\gamma(\xi^n_{X_0}) := \begin{cases} \operatorname{sign}(\gamma)\, \exp(\gamma \xi^n_{X_0}), & \text{if } \gamma \neq 0 \\ \xi^n_{X_0} & \text{for } \gamma = 0. \end{cases} \tag{1.2}$$

Obviously, if $\gamma = 0$ then $u^\gamma(\xi^n_{X_0}(\pi)) = \sum_{k=0}^{n-1} r_{X_k,X_{k+1}}(f^k_{X_k})$.

Supposing that the chain starts in state $X_0 = i$ and policy $\pi = (f^n)$ is followed, then for expected utility in the n next transitions we have (E^π_i denotes expectation if policy π is followed and the starting state $X_0 = i$)

$$\bar{U}^\pi_i(\gamma, 0, n) \;:=\; \mathrm{E}^\pi_i[u^\gamma(\xi^n_{X_0})] = (\operatorname{sign}\,\gamma)\, U^\pi_i(\gamma, 0, n) \tag{1.3}$$

where

$$U^\pi_i(\gamma, 0, n) \;:=\; \mathrm{E}^\pi_i\left[\exp\left(\gamma \sum_{k=0}^{n-1} r_{X_k,X_{k+1}}(f^k_{X_k})\right)\right] > 0. \tag{1.4}$$

Similarly, for $m < n$ if the starting state $X_m = i$ we write

$$U^\pi_i(\gamma, m, n) := \mathrm{E}^\pi_i\left[\exp\left(\gamma \sum_{k=m}^{n-1} r_{X_k,X_{k+1}}(f^k_{X_k})\right)\right]. \tag{1.5}$$

Moreover, let $G^\pi_i(\gamma) \in \mathbb{R}^+$ be the *growth rate of* $U^\pi_i(\gamma, 0, n)$ defined implicitly by

$$\alpha_1\, (G^\pi_i(\gamma))^n \le U^\pi_i(\gamma, 0, n) \le \alpha_2\, (G^\pi_i(\gamma))^n \tag{1.6}$$

where real numbers $\alpha_2 > \alpha_1 > 0$.
In addition, if $\gamma \neq 0$ for the associated certainty equivalent, say $Z^\pi_i(\gamma, 0, n)$, defined implicitly by $u^\gamma(Z^\pi_i(\gamma, 0, n)) := \mathrm{E}^\pi_i[u^\gamma(\xi^n_{X_0}(f^k_{X_k}))]$, and for its asymptotical mean

value, say $J_i^\pi(\gamma,0)$, we have

$$Z_i^\pi(\gamma,0,n) = \frac{1}{\gamma}\ln\left\{\mathrm{E}_i^\pi\left[\exp\left(\gamma\sum_{k=0}^{n-1} r_{X_k,X_{k+1}}(f_{X_k}^k)\right)\right]\right\} \tag{1.7}$$

$$J_i^\pi(\gamma,0) = \limsup_{n\to\infty}\frac{1}{n}Z_i^\pi(\gamma,0,n). \tag{1.8}$$

Note that for $\gamma=0$ we have $Z_i^\pi(\gamma,0,n)=\mathrm{E}_i^\pi[\sum_{k=0}^{n-1} r_{X_k,X_{k+1}}(f_{X_k}^k)]$ (the standard expected reward criterion) and $J_i^\pi(\gamma,0)$ is the corresponding mean value.

In what follows we shall often abbreviate $\bar{U}_i^\pi(\gamma,0,n)$, $U_i^\pi(\gamma,0,n)$, $Z_i^\pi(\gamma,0,n)$ and $J_i^\pi(\gamma,0)$ respectively by $\bar{U}_i^\pi(\gamma,n)$, $U_i^\pi(\gamma,n)$, $Z_i^\pi(\gamma,n)$ and $J_i^\pi(\gamma)$ respectively. Similarly $\boldsymbol{U}^\pi(\gamma,n)$ (resp. $\boldsymbol{Z}^\pi(\gamma,n)$, resp. $\boldsymbol{J}^\pi(\gamma)$) is reserved for the vector of expected utilities (resp. certainty equivalents, resp. mean values of certainty equivalents) whose ith element equals $U_i^\pi(\gamma,n)$ (resp. $Z_i^\pi(\gamma,n)$, resp. $J_i^\pi(\gamma)$). The symbol $\boldsymbol{I}$ is reserved for an identity matrix and $\boldsymbol{e}$ is a unit (column) vector. Moreover, for any $f\in\mathcal{F}$, let

$$\boldsymbol{Q}^{(\gamma)}(f) = \left[q_{ij}^{(\gamma)}(f_i)\right] \tag{1.9}$$

be an $N\times N$ nonnegative matrix with elements

$$q_{ij}^{(\gamma)}(f_i) := p_{ij}(f_i)\exp(\gamma\, r_{ij}(f_i)). \tag{1.10}$$

Observe that $\boldsymbol{Q}^{(\gamma)}(f)$ is irreducible if and only if $\boldsymbol{P}(f)$ is irreducible, and a class of states is closed in $\boldsymbol{Q}^{(\gamma)}(f)$ if and only if it is closed in $\boldsymbol{P}(f)$. Similarly as in the Markov chain theory we can speak of accessibility of elements (states) of the matrix $\boldsymbol{Q}^{(\gamma)}(f)$. Moreover, if $\widetilde{\mathcal{I}}(f)\subset\mathcal{I}$ is a closed set of transient states (with the corresponding submatrix $\widetilde{\boldsymbol{P}}(f)$ having the spectral radius less than unity), then the spectral radius of $\widetilde{\boldsymbol{Q}}^{(\gamma)}(f)$ may be equal to the spectral radius of $\boldsymbol{Q}^{(\gamma)}(f)$ and even greater than the spectral radius of any other irreducible class of $\boldsymbol{Q}^{(\gamma)}(f)$. Finally, observe that, similarly to the "product property" of the set of "transition probability matrices" arising in standard models of dynamic programming, the considered collection $\{\boldsymbol{Q}^{(\gamma)}(f), f\in\mathcal{F}\}$ of nonnegative matrices also fulfills the "product property."

In this note we focus attention on the characterization of policies maximizing growth rate of expected utility, along with average of the associated certainty equivalent. It is known from the literature that for communicating Markov chains (and also for unichain models with the risk-sensitivity close to zero) optimal average values of certainty equivalents are independent of the starting state (see [5, 8]). In contrast to the existing literature our analysis is based on methods of stochastic dynamic programming on condition that the transition probabilities are replaced by general nonnegative matrices. In particular, we focus on the properties of (in general) nonhomogeneous matrix products selected from a collection of nonnegative matrices $\{\boldsymbol{Q}^{(\gamma)}(f), f\in\mathcal{F}\}$ arising in the recursive formulas for the growth of expected utilities. Using the block-triangular decomposition of the set of nonnegative matrices and the existence of some "dominating" matrix in the above mentioned matrix collection, we establish necessary and sufficient conditions guaranteeing independence of the growth rates and average optimal values on starting state along with partition of the state space into subsets with constant growth rates and average optimal values.

Alternatively, instead of transition rewards it is also possible to consider transition costs $c_{ij}(a)$; then we are trying to minimize the considered utility function. In particular, for models with growth rates (and average optimality) independent of the starting state we show how our methods work if we minimize the growth rate (or average of the associated certainty equivalent).

The paper is organized as follows. Section 2 summarizes some useful facts on nonnegative matrices and presents reformulation of the problem in terms of stochastic dynamic programming where transition probability matrices are replaced by general nonnegative matrices. In Section 3 we present necessary and sufficient condition guaranteeing that the growth rate and average of the associated certainty equivalent are independent of the starting conditions. Recalling the uniform block-triangular decomposition of a collection of nonnegative matrices fulfilling the "product property" (see [15, 18, 19, 20]) in Section 4 we are able to decompose the state space in the classes with the same growth rate and same average optimality. In Section 5 we indicate how the obtained results can be employed if we minimize the considered utility function, i. e., instead of transition rewards we consider transition costs. Finally, conclusions and comparison with current results are made in Section 6. In the Appendix we present a slight modification of a policy iteration algorithm (originally suggested in [12]) for finding policies minimizing the growth rate and/or average optimality for models with constant growth rates.

2. CONNECTIONS WITH NONNEGATIVE MATRICES

Since the exponential utility function $u^\gamma(\cdot)$ is separable and the considered control policy π is Markovian, from $u^\gamma(\xi^n_{X_0}) = \exp(\gamma r_{X_0,X_1}) \cdot u^\gamma(\xi^{(1,n)}_{X_1})$ (here for the sake of brevity we omit arguments π and $f^k_{X_k}$) on taking expectations we conclude that (observe that $\mathrm{E}[u^\gamma(r_{X_0,X_1}(f^0_{X_0}))] = \sum_{j\in\mathcal{I}} p_{X_0,j}(f^0_{X_0}) \exp(\gamma\, r_{X_0,j}(f^0_{X_0})))$

$$\begin{aligned} \mathrm{E}^\pi_i u^\gamma(\xi^n_{X_0}) &= \mathrm{E}^\pi_i \{\mathrm{E}\, \mathrm{e}^{r_{i,X_1}(f^0_i)} \cdot \mathrm{E}^\pi_{X_1} u^\gamma(\xi^{(1,n)}_{X_1}) | X_1\} \\ &= \sum_{j\in\mathcal{I}} p_{ij}(f^0_i) \mathrm{e}^{\gamma r_{ij}(f^0_i)} \mathrm{E}^\pi_j u^\gamma(\xi^{(1,n)}_{X_1}) \end{aligned} \tag{2.1}$$

that can be also written as (recall that $q^\gamma_{ij}(\cdot) = p_{ij}(\cdot)\mathrm{e}^{\gamma r_{ij}(\cdot)}$)

$$U^\pi_i(\gamma, 0, n) = \sum_{j\in\mathcal{I}} q^{(\gamma)}_{ij}(f^0_i) \cdot U^\pi_j(\gamma, 1, n) \qquad \text{with } \; U^\pi_i(\gamma, n, n) = 1 \tag{2.2}$$

or in vector notation

$$\boldsymbol{U}^\pi(\gamma, 0, n) = \boldsymbol{Q}^{(\gamma)}(f^0) \cdot \boldsymbol{U}^\pi(\gamma, 1, n) \qquad \text{with } \; \boldsymbol{U}^\pi(\gamma, n, n) = \boldsymbol{e}. \tag{2.3}$$

Iterating (2.3) we get if policy $\pi = (f^n)$ is followed

$$\boldsymbol{U}^{\pi}(\gamma, n) = \boldsymbol{Q}^{(\gamma)}(f^0) \cdot \boldsymbol{Q}^{(\gamma)}(f^1) \cdot \ldots \cdot \boldsymbol{Q}^{(\gamma)}(f^{n-1}) \cdot \boldsymbol{e}. \tag{2.4}$$

In particular, (Markovian) policy $\hat{\pi}^{(n)} = (\hat{f}^{(k,n)})$ maximizing $\boldsymbol{U}^{\pi}(\gamma, n)$, i. e. $\boldsymbol{U}^{\hat{\pi}}(\gamma, n) = \max_{\pi} \boldsymbol{U}^{\pi}(\gamma, n)$ must fulfill the following dynamic programming recursion

$$\begin{aligned} \boldsymbol{U}^{\hat{\pi}}(\gamma, k, n) &= \max_{f \in \mathcal{F}} \{\boldsymbol{Q}^{(\gamma)}(f) \cdot \boldsymbol{U}^{\pi}(\gamma, k+1, n)\} \\ &=: \boldsymbol{Q}^{(\gamma)}(\hat{f}^{(k,n)}) \cdot \boldsymbol{U}^{\pi}(\gamma, k+1, n) \quad \text{for} \quad k = 0, 1, \ldots, n-1 \end{aligned} \tag{2.5}$$

$$\boldsymbol{U}^{\hat{\pi}}(\gamma, n-1, n) = \max_{f \in \mathcal{F}} \{\boldsymbol{Q}^{(\gamma)}(f) \cdot \boldsymbol{e}\} =: \boldsymbol{Q}^{(\gamma)}(\hat{f}^{(n-1,n)}) \cdot \boldsymbol{e}. \tag{2.6}$$

(Here the vectorial maximum is considered componentwise and always exists since the ith row of $\boldsymbol{Q}^{(\gamma)}(f)$ depends only on the decision (action) taken in state i, cf. the "product property" of the set of matrices $\boldsymbol{Q}^{(\gamma)}(f)$'s.)

Since $\boldsymbol{Q}^{(\gamma)}(f)$ is a nonnegative matrix, by the well-known Perron–Frobenius theorem (see, e. g. [3, 11]) the spectral radius $\rho^{(\gamma)}(f)$ of $\boldsymbol{Q}^{(\gamma)}(f)$ is equal to the maximum positive eigenvalue of $\boldsymbol{Q}^{(\gamma)}(f)$ and the corresponding left (row) and right (column) eigenvectors, say $\boldsymbol{y}^{(\gamma)}(f)$, $\boldsymbol{x}^{(\gamma)}(f)$, (called the Perron eigenvectors) can be selected nonnegative. In particular, it holds

$$\rho^{(\gamma)}(f)\, \boldsymbol{y}^{(\gamma)}(f) = \boldsymbol{y}^{(\gamma)}(f) \cdot \boldsymbol{Q}^{(\gamma)}(f) \qquad \text{with} \quad \boldsymbol{y}^{(\gamma)}(f) \geq \boldsymbol{0} \tag{2.7}$$

$$\rho^{(\gamma)}(f)\, \boldsymbol{x}^{(\gamma)}(f) = \boldsymbol{Q}^{(\gamma)}(f) \cdot \boldsymbol{x}^{(\gamma)}(f) \qquad \text{with} \quad \boldsymbol{x}^{(\gamma)}(f) \geq \boldsymbol{0}. \tag{2.8}$$

In case that $\boldsymbol{Q}^{(\gamma)}(f)$ is irreducible (i. e. if $\boldsymbol{P}(f)$ is irreducible) the Perron eigenvectors can be selected strictly positive, i. e. (2.7), (2.8) hold with $\boldsymbol{y}^{(\gamma)}(f) > \boldsymbol{0}$, $\boldsymbol{x}^{(\gamma)}(f) > \boldsymbol{0}$. (In a vector inequality $\boldsymbol{a} \geq \boldsymbol{b}$ denotes that $a_i \geq b_i$ for all elements of the vectors $\boldsymbol{a}$, $\boldsymbol{b}$, and $a_i > b_i$ at least for one i, but not for all i's and $\boldsymbol{a} > \boldsymbol{b}$ if and only if and $a_i > b_i$ for all i's.)

Moreover, strictly positive Perron eigenvectors still exist for reducible nonnegative matrices with a specific structure. Necessary and sufficient condition for the existence of a strictly positive right eigenvector $\boldsymbol{x}^{(\gamma)}(f)$ of a nonnegative matrix $\boldsymbol{Q}^{(\gamma)}(f)$ with $f \in \mathcal{F}$ can be formulated as follows (see, e. g. [3, 11]):

If for suitable labelling of states of the underlying Markov chain (i. e. on suitably permuting rows and corresponding columns of $\boldsymbol{Q}^{(\gamma)}(f)$) it is possible to decompose $\boldsymbol{Q}^{(\gamma)}(f)$ on the following block-triangular form:

$$\boldsymbol{Q}^{(\gamma)}(f) = \begin{bmatrix} \boldsymbol{Q}^{(\gamma)}_{(\mathrm{NN})}(f) & \boldsymbol{Q}^{(\gamma)}_{(\mathrm{NB})}(f) \\ \boldsymbol{0} & \boldsymbol{Q}^{(\gamma)}_{(\mathrm{BB})}(f) \end{bmatrix} \tag{2.9}$$

where $\boldsymbol{Q}^{(\gamma)}_{(\mathrm{NN})}(f)$ and $\boldsymbol{Q}^{(\gamma)}_{(\mathrm{BB})}(f)$ (with spectral radius $\rho^{(\gamma)}_{(\mathrm{N})}(f)$ and $\rho^{(\gamma)}_{(\mathrm{B})}(f)$) are (in general reducible) matrices such that:

- $\rho^{(\gamma)}_{(\mathrm{N})}(f) < \rho^{(\gamma)}(f)$,

- $\rho^{(\gamma)}_{(\mathrm{B})}(f) = \rho^{(\gamma)}(f)$ and $\boldsymbol{Q}^{(\gamma)}_{(\mathrm{BB})}(f)$ is diagonal, in particular,

$$\boldsymbol{Q}^{(\gamma)}_{(\mathrm{BB})}(f) = \begin{bmatrix} \boldsymbol{Q}^{(\gamma)}_{(11)}(f) & \dots & \mathbf{0} \\ \vdots & \ddots & \vdots \\ \mathbf{0} & \dots & \boldsymbol{Q}^{(\gamma)}_{(rr)}(f) \end{bmatrix} \tag{2.10}$$

  where $\boldsymbol{Q}^{(\gamma)}_{(ii)}(f)$ (with $i = 1, \dots, r$) are irreducible submatrices (the so-called *basic classes* of $\boldsymbol{Q}^{(\gamma)}(f)$) such that the spectral radius $\rho^{(\gamma)}_i(f)$ of every $\boldsymbol{Q}^{(\gamma)}_{(ii)}(f)$ (with $i = 1, \dots, r$) is equal to $\rho^{(\gamma)}(f)$,

- each irreducible class of $\boldsymbol{Q}^{(\gamma)}_{(\mathrm{NN})}(f)$ (such a class is a *non-basic class* of $\boldsymbol{Q}^{(\gamma)}(f)$, i. e., its spectral radius is less than $\rho^{(\gamma)}(f)$) has access to some basic class of $\boldsymbol{Q}^{(\gamma)}(f)$ (accessibility is considered with respect to the underlying Markov chain $\boldsymbol{P}(f)$, hence at least some elements of $\boldsymbol{Q}^{(\gamma)}_{(\mathrm{NB})}(f)$ are nonvanishing).

Observe that (2.8), (2.9) well correspond to the canonical decomposition of a multichain transition probability matrix.

Remark. Here and in the sequel subscript N (in roman) is reserved for non-basic classes and subscript B (in roman) is reserved for basic classes; on the contrary to latin N reserved for the dimension of the state space $\mathcal{I}$.

Moreover, under condition that $\boldsymbol{x}^{(\gamma)}(f) > \mathbf{0}$ for each $f \in \mathcal{F}$, it can be shown that there exists decision vector $\hat{f} \in \mathcal{F}$ such that $\rho^{(\gamma)}(\hat{f}) \equiv \hat{\rho}^{(\gamma)}$ is the maximum possible eigenvalue of $\boldsymbol{Q}^{(\gamma)}(f)$ over all $f \in \mathcal{F}$, and

$$\begin{aligned} \boldsymbol{Q}^{(\gamma)}(f) \cdot \boldsymbol{x}^{(\gamma)}(\hat{f}) \ &\le \ \max_{f \in \mathcal{F}} \{\boldsymbol{Q}^{(\gamma)}(f) \cdot \boldsymbol{x}^{(\gamma)}(\hat{f})\} \\ = \boldsymbol{Q}^{(\gamma)}(\hat{f}) \cdot \boldsymbol{x}^{(\gamma)}(\hat{f}) \ &= \ \rho^{(\gamma)}(\hat{f}) \boldsymbol{x}^{(\gamma)}(\hat{f}), \quad \text{with} \quad \boldsymbol{x}^{(\gamma)}(\hat{f}) > \mathbf{0}. \end{aligned} \tag{2.11}$$

If all $\boldsymbol{Q}^{(\gamma)}(f)$'s are irreducible, the (constructive) proof based on policy iterations can be found in [12], its extension to (reducible) matrices having strictly positive right eigenvectors can be found in the Appendix. For further extensions to the case of general reducible matrices see [15, 17, 18, 19, 20].

3. MODELS WITH CONSTANT GROWTH RATES AND CONSTANT AVERAGE CERTAINTY EQUIVALENTS

In this section we consider risk-sensitive Markov decision chains where the maximal growth rate (or equivalently mean values of the certainty equivalent) is independent of the starting condition. In contrast to the existing literature our approach is based on the analysis of the growth rate of (in general nonhomogeneous) products of a family of nonnegative matrices and relations between the growth rate and certainty equivalents, cf. (1.4) – (1.8).

Recalling (2.11) we make the following assumption.

Assumption 3.1. There exists $\hat{\rho}^{(\gamma)} \equiv \rho^{(\gamma)}(\hat{f})$ and $\widehat{\boldsymbol{x}}^{(\gamma)} \equiv \boldsymbol{x}^{(\gamma)}(\hat{f}) > \boldsymbol{0}$ (unique up to a multiplicative constant) such that (for a given value of the risk aversion coefficient γ)

$$\hat{\rho}^{(\gamma)}\, \widehat{\boldsymbol{x}}^{(\gamma)} = \max_{f \in \mathcal{F}} \{\boldsymbol{Q}^{(\gamma)}(f) \cdot \widehat{\boldsymbol{x}}^{(\gamma)}\} = \boldsymbol{Q}^{(\gamma)}(\hat{f}) \cdot \widehat{\boldsymbol{x}}^{(\gamma)} \tag{3.1}$$

such that on using the matrix decomposition according to (2.8), i. e., on writing

$$\boldsymbol{Q}^{(\gamma)}(\hat{f}) = \begin{bmatrix} \boldsymbol{Q}^{(\gamma)}_{(\mathrm{NN})}(\hat{f}) & \boldsymbol{Q}^{(\gamma)}_{(\mathrm{NB})}(\hat{f}) \\ \boldsymbol{0} & \boldsymbol{Q}^{(\gamma)}_{(\mathrm{BB})}(\hat{f}) \end{bmatrix} \tag{3.2}$$

where $\boldsymbol{Q}^{(\gamma)}_{(\mathrm{BB})}(\hat{f})$ is the "biggest" diagonal class with spectral radius $\rho^{(\gamma)}(\hat{f})$ among all $\boldsymbol{Q}^{(\gamma)}(f)$'s fulfilling (3.1).

Theorem 3.1. If condition (3.1) holds then for a given γ there exist numbers $\alpha_2^{(\gamma)} > \alpha_1^{(\gamma)} > 0$ such that

$$\alpha_1^{(\gamma)}\, \boldsymbol{x}^{(\gamma)}(\hat{f}) \le (\hat{\rho}^{(\gamma)})^{-n} \prod_{k=0}^{n-1} \boldsymbol{Q}^{(\gamma)}(f^k) \cdot \boldsymbol{e} \le \alpha_2^{(\gamma)}\, \boldsymbol{x}^{(\gamma)}(\hat{f}) \tag{3.3}$$

for any policy $\pi = (f^k)$ maximizing the growth of $\boldsymbol{U}^\pi(\gamma, n)$ for $n = 0, 1, \ldots$. In addition, (3.3) is also fulfilled for stationary policy $\hat{\pi} \sim (\hat{f})$.

Proof. If Assumption 3.1 holds (with $\boldsymbol{x}^{(\gamma)}(f)$ not necessarily strictly positive for each $f \in \mathcal{F}$), we can select $\boldsymbol{x}^{(\gamma)}(\hat{f}) > \boldsymbol{0}$ such that either $\boldsymbol{x}^{(\gamma)}(\hat{f}) \ge \boldsymbol{e}$ or $\boldsymbol{x}^{(\gamma)}(\hat{f}) \le \boldsymbol{e}$. Iterating (2.11) we can immediately conclude that for $\boldsymbol{x}^{(\gamma)}(\hat{f}) \ge \boldsymbol{e}$ and any policy $\pi = (f^k)$

$$\begin{aligned} \prod_{k=0}^{n-1} \boldsymbol{Q}^{(\gamma)}(f^k) \cdot \boldsymbol{e} &\le \prod_{k=0}^{n-1} \boldsymbol{Q}^{(\gamma)}(f^k) \cdot \boldsymbol{x}^{(\gamma)}(\hat{f}) \\ &\le (\boldsymbol{Q}^{(\gamma)}(\hat{f}))^n \cdot \boldsymbol{x}^{(\gamma)}(\hat{f}) = (\hat{\rho}^{(\gamma)})^n\, \boldsymbol{x}^{(\gamma)}(\hat{f}) \end{aligned} \tag{3.4}$$

and hence the asymptotic behaviour of $\boldsymbol{U}^\pi(\gamma, n)$ (or of $\boldsymbol{U}^\pi(\gamma, m, n)$ if m is fixed) heavily depends on $\rho^{(\gamma)}(\hat{f}) \equiv \hat{\rho}^{(\gamma)}$, and elements of $\prod_{k=0}^{n-1} \boldsymbol{Q}^{(\gamma)}(f^k) \cdot \boldsymbol{x}^{(\gamma)}(\hat{f})$ must be bounded from above by $(\hat{\rho^{(\gamma)}})^n \cdot \boldsymbol{x}^{(\gamma)}(\hat{f})$.

Similarly, on selecting $\boldsymbol{x}^{(\gamma)}(\hat{f}) \le \boldsymbol{e}$ from (2.5), (2.11) we get for any policy $\hat{\pi}^{(n)} = (\hat{f}^{(k,n)})$ maximizing $\boldsymbol{U}^{\pi}(\gamma, n)$:

$$\prod_{k=0}^{n-1} \boldsymbol{Q}^{(\gamma)}(\hat{f}^{(k,n)}) \cdot \boldsymbol{e} \ge \prod_{k=0}^{n-1} \boldsymbol{Q}^{(\gamma)}(\hat{f}) \cdot \boldsymbol{e} \ge (\boldsymbol{Q}^{(\gamma)}(\hat{f}))^n \cdot \boldsymbol{x}^{(\gamma)}(\hat{f}) = (\hat{\rho}^{(\gamma)})^n \cdot \boldsymbol{x}^{(\gamma)}(\hat{f}). \quad (3.5)$$

Hence the growth of $\boldsymbol{U}^{\pi}(\gamma, n)$ if a policy maximizing $\boldsymbol{U}^{\pi}(\gamma, n)$ is followed is bounded from below by $(\rho^{\hat{(\gamma)}})^n \cdot \boldsymbol{x}^{(\gamma)}(\hat{f})$.

From (3.4) and (3.5) we immediately get conclusions of Theorem 3.1. □

For what follows it is convenient to rephrase Theorem 3.1 in words as

Corollary 3.2. Under condition (3.1) if policy $\pi = (f^n)$ maximizing $\boldsymbol{U}^{\pi}(\gamma, n)$ is followed the *growth rate* of each element of $\boldsymbol{U}^{\pi}(\gamma, n) = \prod_{k=0}^{n-1} \boldsymbol{Q}^{(\gamma)}(f^k) \cdot \boldsymbol{e}$ is the same and equals $\hat{\rho}^{(\gamma)}$. Moreover, stationary policy $\hat{\pi} \sim (\hat{f})$ also maximizes the growth rate.

Denoting elements of $\boldsymbol{x}^{(\gamma)}(\hat{f}) > \boldsymbol{0}$ by $x_j^{(\gamma)}(f)$ (for $j = 1, \dots, N$) and elements of an $N \times N$ matrix $\boldsymbol{Q}^{(\gamma)}(f)$ by $q_{ij}^{(\gamma)}(f_i)$ (recall that by (1.10) $q_{ij}^{(\gamma)}(f_i) = p_{ij}(f_i)\, \mathrm{e}^{\gamma\, r_{ij}(f_i)}$), for

$g(f) := \gamma^{-1} \ln(\rho^{(\gamma)}(f)), \qquad w_j(f) := \gamma^{-1} \ln(x_j^{(\gamma)}(f))$ (with $j = 1, \dots, N$)

(2.11), (3.1) can be also written as the following set of (nonlinear) equations:

$$\mathrm{e}^{\gamma\,(g(\hat{f}) + w_i(\hat{f}))} = \max_{a \in \mathcal{A}_i} \left\{ \sum_{j \in \mathcal{I}} p_{ij}(a) \cdot \mathrm{e}^{\gamma\,(r_{ij}(a) + w_j(\hat{f}))} \right\}, \quad \text{for} \quad i = 1, \dots, N, \quad (3.6)$$

called γ-*average reward optimality equation.*

In the multiplicative form (used before) (3.6) takes on the form:

$$\rho^{(\gamma)}(\hat{f})\, x_i^{(\gamma)}(\hat{f}) = \max_{a \in \mathcal{A}_i} \left\{ \sum_{j \in \mathcal{I}} p_{ij}(a) \cdot \mathrm{e}^{\gamma r_{ij}(a)} \cdot x_j^{(\gamma)}(\hat{f}) \right\}, \quad \text{for} \quad i = 1, \dots, N. \quad (3.7)$$

Observe that the solution to (3.6), resp. (3.7), i. e. $g(\hat{f})$, $w_i(\hat{f})$, resp. $\rho^{(\gamma)}(\hat{f})$, $x_i(\hat{f})$ is unique up to an additive constant (added to $w_i(\hat{f})$'s), resp. multiplicative constant (applied to $x_i^{(\gamma)}(\hat{f})$'s) and the matrix $\boldsymbol{P}(f) = [p_{ij}(f)]$ occurring in (3.6), (3.7) may be periodic.

Using the above facts Corollary 3.2 can be formulated as

Theorem 3.3. If condition (3.1) holds then for any policy $\pi = (f^n)$ the asymptotical mean value $J_i^{\pi}(\gamma, 0)$ is bounded from above by $g(\hat{f}) := \gamma^{-1} \ln(\rho^{(\gamma)}(\hat{f}))$. Moreover, stationary policy $\hat{\pi} \sim (\hat{f})$ yields the maximum asymptotical mean value $J_i^{\pi}(\gamma, 0)$ that is independent of the starting state $i \in \mathcal{I}$ and equal to $g(\hat{f}) := \gamma^{-1} \ln(\rho^{(\gamma)}(\hat{f}))$.

Assumption 3.2. For each $f \in \mathcal{F}$ the transition probability matrix $\boldsymbol{P}(f)$ has a single recurrent class, i. e., we assume existence of some state, say N, that is accessible from any state $i \in \mathcal{I}$ under each $f \in \mathcal{F}$ (hence N is recurrent under each $f \in \mathcal{F}$).

Under Assumption 3.2 for suitable labelling of states (i. e., on suitably permuting rows and corresponding columns of $\boldsymbol{P}(f)$) it is possible to decompose $\boldsymbol{P}(f)$ such that:

$$\boldsymbol{P}(f) = \begin{bmatrix} \boldsymbol{P}_{(\mathrm{NN})}(f) & \boldsymbol{P}_{(\mathrm{NB})}(f) \\ \boldsymbol{0} & \boldsymbol{P}_{(\mathrm{BB})}(f) \end{bmatrix} \tag{3.8}$$

where (in general reducible) submatrix $\boldsymbol{P}_{(\mathrm{NN})}(f)$ contains all transient states of $\boldsymbol{P}(f)$ and an irreducible class $\boldsymbol{P}_{(\mathrm{BB})}(f)$ containing all recurrent states is a single submatrix of $\boldsymbol{P}(f)$ with spectral radius equal to one. Of course, the substochastic matrix $\boldsymbol{P}_{(\mathrm{NN})}(f)$ can be further decomposed in the following block-triangular form ($\boldsymbol{P}_{(\mathrm{N})ii}(f)$ are irreducible classes of transient states)

$$\boldsymbol{P}_{(\mathrm{NN})}(f) = \begin{bmatrix} \boldsymbol{P}_{(\mathrm{N})11}(f) & \boldsymbol{P}_{(\mathrm{N})12}(f) & \dots & \boldsymbol{P}_{(\mathrm{N})1u}(f) \\ \boldsymbol{0} & \boldsymbol{P}_{(\mathrm{N})22}(f) & \dots & \boldsymbol{P}_{(\mathrm{N})2u}(f) \\ \vdots & \vdots & \ddots & \vdots \\ \boldsymbol{0} & \boldsymbol{0} & \dots & \boldsymbol{P}_{(\mathrm{N})uu}(f) \end{bmatrix}. \tag{3.9}$$

Since the matrix $\boldsymbol{Q}^{(\gamma)}(f)$ is generated from $\boldsymbol{P}(f)$ by multiplying its ijth entry by $\mathrm{e}^{\gamma\, r_{ij}(f_i)}$ and since its spectral radius $\rho^{(\gamma)}(f)$ is a continuous function of the matrix elements[1], the matrix $\boldsymbol{Q}^{(\gamma)}(f)$ still have a strictly positive right (Perron) eigenvector if the risk aversion coefficient γ is sufficiently close to null. Of course, if at least for one pair of transient states, say i_0, j_0, belonging to the same irreducible class it holds $r_{i_0,j_0}(\cdot) > r_{ij}(\cdot)$ for any pair of states i, j belonging the recurrent class $\boldsymbol{P}_{(\mathrm{BB})}(f)$ of $\boldsymbol{P}(f)$ then for sufficiently large risk-aversion coefficient γ it happens that the class $\boldsymbol{P}_{(\mathrm{BB})}(f)$ of recurrent state is no more the basic class of the corresponding matrix $\boldsymbol{Q}^{(\gamma)}(f)$ and hence there exists no strictly positive right Perron eigenvector of the matrix $\boldsymbol{Q}^{(\gamma)}(f)$ as the following example can show.

Example 1. Consider an uncontrolled model (hence the argument f is deleted) where $N = 4$; $p_{ij} = 0.25$, $r_{ij} = 1$, for $i, j = 1, 2$; $p_{ij} = 0.25$, $r_{ij} = 0$, for $i = 1, 2$, $j = 3, 4$; and for $i = 3, 4$ we have $p_{i1} = p_{i2} = 0$, $r_{i1} = r_{i2} = 0$, $p_{i3} = p_{i4} = 0.5$, $r_{i3} = r_{i4} = 0$. Hence using the decomposition according to (3.8) we have

$$\boldsymbol{P}_{(\mathrm{NN})} = \begin{bmatrix} 0.25 & 0.25 \\ 0.25 & 0.25 \end{bmatrix}, \quad \boldsymbol{P}_{(\mathrm{NB})} = \begin{bmatrix} 0.25 & 0.25 \\ 0.25 & 0.25 \end{bmatrix}, \quad \boldsymbol{P}_{(\mathrm{BB})} = \begin{bmatrix} 0.5 & 0.5 \\ 0.5 & 0.5 \end{bmatrix}$$

where

$$\boldsymbol{P}_{(\mathrm{NB})} = \boldsymbol{Q}^{(\gamma)}_{(\mathrm{NB})},\ \boldsymbol{P}_{(\mathrm{BB})} = \boldsymbol{Q}^{(\gamma)}_{(\mathrm{BB})} \text{ for any } \gamma \in \mathbb{R}, \text{ but } \quad \boldsymbol{Q}^{(\gamma)}_{(\mathrm{NN})} = \mathrm{e}^{\gamma}\boldsymbol{P}_{(\mathrm{NN})}.$$

[1]Recall that eigenvalues of a (finite) dimensional matrix can be calculated as the roots of the respective a characteristic polynomial being a continuous function of the matrix elements.

Then the spectral radius $\rho^{(\gamma)}_{(\mathrm{B})}$ of $\boldsymbol{Q}^{(\gamma)}_{(\mathrm{BB})} = \boldsymbol{P}_{(\mathrm{BB})}$ equals 1, and the spectral radius $\rho^{(\gamma)}_{(\mathrm{N})}$ of $\boldsymbol{Q}^{(\gamma)}_{(\mathrm{NN})}$ is equal to $0.5\,\mathrm{e}^{\gamma}$. Obviously, the corresponding (right) eigenvectors of $\boldsymbol{Q}^{(\gamma)}_{(\mathrm{BB})}$ and $\boldsymbol{Q}^{(\gamma)}_{(\mathrm{NN})}$ are two-dimensional unit vectors. Observe that for $\gamma = \ln 2$ $\rho^{(\gamma)}_{(\mathrm{N})} = \rho^{(\gamma)}_{(\mathrm{B})} = 1$.

Hence if $\gamma < \ln 2$, for the spectral radius of $\boldsymbol{Q}^{(\gamma)}$ we have $\rho^{(\gamma)} = \rho^{(\gamma)}_{(\mathrm{B})} = 1$ and the corresponding right Perron eigenvector $\boldsymbol{x}^{(\gamma)} = [(2-\mathrm{e}^{\gamma})^{-1} \ \ (2-\mathrm{e}^{\gamma})^{-1} \ \ 1 \ \ 1]^{\mathrm{T}}$ is strictly positive.
On the contrary if $\gamma > \ln 2$, for the spectral radius $\rho^{(\gamma)}$ of $\boldsymbol{Q}^{(\gamma)}$ we have $\rho^{(\gamma)} = 0.5\mathrm{e}^{\gamma} > 1$ and the corresponding right Perron eigenvector $\boldsymbol{x}^{(\gamma)} = [1 \ \ 1 \ \ 0 \ \ 0]^{\mathrm{T}}$ is not strictly positive.
Conclusions:
If $\gamma < \ln 2$ then the the growth rate $G^{\pi}_i(\gamma) = 1$ and the asymptotic mean values $J^{\pi}_i(\gamma) = 0$ are independent of the starting state i. On the other hand:
If $\gamma > \ln 2$ the growth rate $G^{\pi}_i(\gamma) = 1$ only for $i = 3, 4$, but $G^{\pi}_i(\gamma) = \mathrm{e}^{\gamma}$ for $i = 1, 2$.

4. MODELS WITH NON–CONSTANT GROWTH RATES AND NON–CONSTANT AVERAGE CERTAINTY EQUIVALENTS

In this section we consider risk-sensitive models with non-constant growth rates and non-constant average values of the certainty equivalents. To this end at least some matrices pertaining to the set $\{\boldsymbol{Q}^{(\gamma)}(f), f \in \mathcal{F}\}$ must be reducible with no strictly positive right eigenvector.

First observe that for any (reducible) nonnegative matrix $\boldsymbol{Q}^{(\gamma)}(f)$ we can easily identify its basic classes. In case that there exists a single basic class of $\boldsymbol{Q}^{(\gamma)}(f)$, for suitable labelling the states of the underlying Markov chain X or equivalently on suitably permuting rows and corresponding columns of the matrix $\boldsymbol{Q}^{(\gamma)}(f)$, then $\boldsymbol{Q}^{(\gamma)}(f)$ can be decomposed as

$$\boldsymbol{Q}^{(\gamma)}(f) = \begin{bmatrix} \boldsymbol{Q}^{(\gamma)}_{11}(f) & \boldsymbol{Q}^{(\gamma,1)}_{1}(f) \\ \boldsymbol{0} & \boldsymbol{Q}^{(\gamma,1)}(f) \end{bmatrix} \tag{4.1}$$

where the structure of the diagonal block $\boldsymbol{Q}^{(\gamma)}_{11}(f)$ with spectral radius $\rho^{(\gamma)}_1(f) = \rho^{(\gamma)}(f)$ is the same as in (2.9) (i. e. all elements of $\boldsymbol{Q}^{(\gamma)}_{11}(f)$ have access to the basic class of $\boldsymbol{Q}^{(\gamma)}_{11}(f)$), and for the spectral radius of $\boldsymbol{Q}^{(\gamma,1)}(f)$ we have $\rho^{(\gamma,1)}(f) \le \rho^{(\gamma)}(f)$. In particular, in virtue of (2.8), (2.9) we can conclude that

$$\boldsymbol{Q}^{(\gamma)}_{11}(f) = \begin{bmatrix} \boldsymbol{Q}^{(\gamma)}_{1(\mathrm{NN})}(f) & \boldsymbol{Q}^{(\gamma)}_{1(\mathrm{NB})}(f) \\ \boldsymbol{0} & \boldsymbol{Q}^{(\gamma)}_{1(\mathrm{BB})}(f) \end{bmatrix} \tag{4.2}$$

with $\boldsymbol{Q}^{(\gamma)}_{1(\mathrm{BB})}(f)$ being the basic class of $\boldsymbol{Q}^{(\gamma)}_{11}(f)$ (and also of $\boldsymbol{Q}^{(\gamma)}(f)$). Since all elements of $\boldsymbol{Q}^{(\gamma)}_{1(\mathrm{NN})}(f)$ are accessible to $\boldsymbol{Q}^{(\gamma)}_{1(\mathrm{BB})}(f)$, there exists a strictly positive (right) Perron eigenvector of $\boldsymbol{Q}_{11}^{(\gamma)}(f)$, i. e.

$$\boldsymbol{Q}_{11}^{(\gamma)}(f) \cdot \boldsymbol{x}_1^{(\gamma)}(f) = \rho_1^{(\gamma)}(f)\, \boldsymbol{x}_1^{(\gamma)}(f) \tag{4.3}$$

where $\rho_1^{(\gamma)}(f)$ is the spectral radius of $\boldsymbol{Q}_{11}^{(\gamma)}(f)$ (and hence also of $\boldsymbol{Q}^{(\gamma)}(f)$) and $\boldsymbol{x}_1^{(\gamma)}(f) > \mathbf{0}$ is the corresponding (right) Perron eigenvector.

Similarly for the diagonal block $\boldsymbol{Q}^{(\gamma,1)}(f)$ (assuming the existence of a single basic class of $\boldsymbol{Q}^{(\gamma,1)}(f)$) on suitably permuting rows and corresponding columns formulas analogous to (4.1)–(4.3) will hold for the matrix $\boldsymbol{Q}^{(\gamma,1)}(f)$ and its upper diagonal block $\boldsymbol{Q}_{22}^{(\gamma)}(f)$

$$\boldsymbol{Q}^{(\gamma,1)}(f) = \begin{bmatrix} \boldsymbol{Q}_{22}^{(\gamma)}(f) & \boldsymbol{Q}_{2}^{(\gamma,2)}(f) \\ \mathbf{0} & \boldsymbol{Q}^{(\gamma,2)}(f) \end{bmatrix}, \qquad \boldsymbol{Q}_{22}^{(\gamma)}(f) = \begin{bmatrix} \boldsymbol{Q}_{2(\mathrm{NN})}^{(\gamma)}(f) & \boldsymbol{Q}_{2(\mathrm{NB})}^{(\gamma)}(f) \\ \mathbf{0} & \boldsymbol{Q}_{2(\mathrm{BB})}^{(\gamma)}(f) \end{bmatrix}$$

along with the diagonal blocks of $\boldsymbol{Q}_{22}^{(\gamma)}(f)$ denoted $\boldsymbol{Q}_{2(\mathrm{NN})}^{(\gamma)}(f)$, $\boldsymbol{Q}_{2(\mathrm{BB})}^{(\gamma)}(f)$ and its "transition" off-diagonal block $\boldsymbol{Q}_{2(\mathrm{NB})}^{(\gamma)}(f)$.

Repeating this reasoning we can conclude that for suitable labelling of states of the underlying Markov chain, or equivalently on suitably permuting rows and corresponding columns, the matrix $\boldsymbol{Q}^{(\gamma)}(f)$ can be decomposed into the following block-triangular form (the number s of diagonal blocks depends on f)

$$\boldsymbol{Q}^{(\gamma)}(f) = \begin{bmatrix} \boldsymbol{Q}_{11}^{(\gamma)}(f) & \boldsymbol{Q}_{12}^{(\gamma)}(f) & \ldots & \boldsymbol{Q}_{1s}^{(\gamma)}(f) \\ \mathbf{0} & \boldsymbol{Q}_{22}^{(\gamma)}(f) & \ldots & \boldsymbol{Q}_{2s}^{(\gamma)}(f) \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \ldots & \boldsymbol{Q}_{ss}^{(\gamma)}(f) \end{bmatrix} \tag{4.4}$$

where $\boldsymbol{Q}_{ii}^{(\gamma)}(f)$'s are the "biggest" submatrices of $\boldsymbol{Q}^{(\gamma)}(f)$ having strictly positive right Perron eigenvectors, i. e., there exist $\boldsymbol{x}_i^{(\gamma)}(f) > \mathbf{0}$ such that for all $i = 1, 2, \ldots, s$

$$\boldsymbol{Q}_{ii}^{(\gamma)}(f) \cdot \boldsymbol{x}_i^{(\gamma)}(f) = \rho_i^{(\gamma)}(f)\, \boldsymbol{x}_i^{(\gamma)}(f) \quad \text{with} \quad \rho_i^{(\gamma)}(f) \ge \rho_{i+1}^{(\gamma)}(f). \tag{4.5}$$

Furthermore, the above results can be extended to the whole collection of non-negative matrices $\{\boldsymbol{Q}^{(\gamma)}(f), f \in \mathcal{F}\}$ as it is summarized in the following theorem (for the proofs see [17, 18, 19, 20]).

Theorem 4.1. There exists $\hat{f} \in \mathcal{F}$ and a suitable labelling of states inducing the partition of the state space, say $\hat{\mathcal{I}} \equiv \bigcup_{i=1}^{s} \mathcal{I}_i(\hat{f})$, called the *basic partition*, such that: Keeping the partition in accordance of $\hat{\mathcal{I}}$ then every $\boldsymbol{Q}^{(\gamma)}(f)$ is block triangular, i. e.

$$\boldsymbol{Q}^{(\gamma)}(f) = \begin{bmatrix} \boldsymbol{Q}_{11}^{(\gamma)}(f) & \boldsymbol{Q}_{12}^{(\gamma)}(f) & \ldots & \boldsymbol{Q}_{1s}^{(\gamma)}(f) \\ \mathbf{0} & \boldsymbol{Q}_{22}^{(\gamma)}(f) & \ldots & \boldsymbol{Q}_{2s}^{(\gamma)}(f) \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \ldots & \boldsymbol{Q}_{ss}^{(\gamma)}(f) \end{bmatrix}, \qquad \forall f \in \mathcal{F} \tag{4.6}$$

where all $\boldsymbol{Q}_{ii}^{(\gamma)}(f)$ have fixed dimensions equal to card $\mathcal{I}_i(\hat{f})$, and for $i = 1,\dots,s$ $\boldsymbol{Q}_{ii}^{(\gamma)}(\hat{f})$'s are the "biggest" submatrices of $\boldsymbol{Q}^{(\gamma)}(f)$ having strictly positive right eigenvectors corresponding to the maximum possible spectral radii of the corresponding submatrices, i.e., there exists $\boldsymbol{Q}_{ii}^{(\gamma)}(\hat{f})$ along with $\boldsymbol{x}_i^{(\gamma)}(\hat{f}) > \mathbf{0}$ such that for any $f \in \mathcal{F}$ and all $i = 1,2,\dots,s$

$$\rho_i^{(\gamma)}(\hat{f}) \ge \rho_i^{(\gamma)}(f); \qquad \rho_i^{(\gamma)}(\hat{f}) \ge \rho_{i+1}^{(\gamma)}(\hat{f}) \tag{4.7}$$

$$\boldsymbol{Q}_{ii}^{(\gamma)}(f)\cdot\boldsymbol{x}_i^{(\gamma)}(\hat{f}) \;\le\; \boldsymbol{Q}_{ii}^{(\gamma)}(\hat{f})\cdot\boldsymbol{x}_i^{(\gamma)}(\hat{f}) = \rho_i^{(\gamma)}(\hat{f})\,\boldsymbol{x}_i^{(\gamma)}(\hat{f}). \tag{4.8}$$

Observe that $\rho_1^{(\gamma)}(\hat{f}) = \rho^{(\gamma)}(\hat{f})$ and that each diagonal block $\boldsymbol{Q}_{ii}^{(\gamma)}(f)$ in (4.6) may be reducible, and if $\boldsymbol{Q}_{ii}^{(\gamma)}(\hat{f})$ is reducible then it can be decomposed similarly as in (4.2).

Throughout this note we make the following assumption:

Assumption 4.1. For a given value of the risk aversion coefficient γ a strict inequalities holds in the second part of (4.7), i.e.:

$$\rho_1^{(\gamma)}(\hat{f}) > \rho_2^{(\gamma)}(\hat{f}) > \ldots > \rho_s^{(\gamma)}(\hat{f}). \tag{4.9}$$

Remark. Observe that the case $\rho_i^{(\gamma)}(f) = \rho_{i+1}^{(\gamma)}(f)$ can be easily excluded, since, if necessary, we may assume that after small perturbations of some values $p_{ij}(f_i)$ and $r_{ij}(f_i)$ (i.e. the perturbation of $q_{ij}^{(\gamma)}(f_i)$), we arrive at $\rho_i^{(\gamma)}(f) > \rho_{i+1}^{(\gamma)}(f)$ and condition (4.9) will be fulfilled (recall that the value of the spectral radius is a continuous function of the matrix elements).

In case that $s = 1$ we have $\boldsymbol{x}^{(\gamma)}(\hat{f}) > \mathbf{0}$. Then by Theorem 3.1 maximum growth rate of $\boldsymbol{U}^{\pi}(\gamma,n)$ is given by $\rho^{(\gamma)}(\hat{f})$, is independent of the starting state and can be obtained if stationary policy $\hat{\pi} \sim (\hat{f})$ is followed.

In case that $s > 1$ by (4.6) we can immediately conclude that on following stationary policy $\hat{\pi} \sim (\hat{f})$ and keeping the basic partition in accordance of Theorem 4.1, if $\boldsymbol{x}_i^{(\gamma)}(\hat{f})$ is selected such that $\boldsymbol{x}_i^{(\gamma)}(\hat{f}) \le \boldsymbol{e}$, then for any $i = 1,2,\dots,s$

$$\boldsymbol{U}_i^{\hat{\pi}}(\gamma,n) = (\boldsymbol{Q}_{ii}^{(\gamma)}(\hat{f}))^n\cdot\boldsymbol{e} \ge (\boldsymbol{Q}_{ii}^{(\gamma)}(\hat{f}))^n\cdot\boldsymbol{x}_i^{(\gamma)}(\hat{f}) \ge (\rho_i^{(\gamma)}(\hat{f}))^n\boldsymbol{x}_i^{(\gamma)}(\hat{f}).$$

In words: For stationary policy $\hat{\pi} \sim (\hat{f})$ the growth rate of every $\boldsymbol{U}_i^{\hat{\pi}}(\gamma,n)$ is nonsmaller than $\rho_i^{(\gamma)}(\hat{f})$.

Hence to establish that the maximal growth rate of elements pertaining to $\mathcal{I}_i(\hat{f})$ equals $\rho_i^{(\gamma)}(\hat{f})$ it is sufficient to show that $\rho_i^{(\gamma)}(\hat{f})$ is also an upper bound on the growth rate of elements from $\mathcal{I}_i(\hat{f})$. To this end, on considering the basic partition, policy π^* generating the maximal growth must fulfil the dynamic programming recursion (2.5) that for the considered reducible case can be also written as:

$$\begin{bmatrix} U_1^*(\gamma,k,n) \\ U_2^*(\gamma,k,n) \\ \vdots \\ U_s^*(\gamma,k,n) \end{bmatrix} = \max_{f\in\mathcal{F}} \begin{bmatrix} Q_{11}^{(\gamma)}(f) & Q_{12}^{(\gamma)}(f) & \dots & Q_{1s}^{(\gamma)}(f) \\ \mathbf{0} & Q_{22}^{(\gamma)}(f) & \dots & Q_{2s}^{(\gamma)}(f) \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \dots & Q_{ss}^{(\gamma)}(f) \end{bmatrix} \cdot \begin{bmatrix} U_1^*(\gamma,k+1,n) \\ U_2^*(\gamma,k+1,n) \\ \vdots \\ U_s^*(\gamma,k+1,n) \end{bmatrix} \tag{4.10}$$

We show by induction on $i = s, s-1, \dots, 1$ that for any $n = 0, 1, \dots,$ the maximal possible growth of each $U_i^*(\gamma,k,n)$ is also dominated by the powers of $\rho_i^{(\gamma)}(\hat f)$.

If $i = s$ the maximal growth rate is given by $Q_{ss}^{(\gamma)}(\hat f)$ by Theorem 3.1 (cf. (3.3)). Hence it suffices only to construct the induction step, i. e., to show that supposing the maximal possible growth of $U_{i+1}^*(\gamma,k,n)$ is dominated by the growth rate equal to $\rho_{i+1}^{(\gamma)}(\hat f)$, then the maximal possible growth of $U_i^*(\gamma,k,n)$ is dominated by $\rho_i^{(\gamma)}(\hat f)$ (where $\rho_i^{(\gamma)}(\hat f) > \rho_{i+1}^{(\gamma)}(\hat f)$ by Assumption 4.1).

For the sake of simplicity we construct the induction step if $i = 1$. To this end let $U^*(\gamma,k,n)$, and $Q^{(\gamma)}(f)$ be decomposed as (cf. (4.1))

$$U^*(\gamma,k,n) = \begin{bmatrix} U_1^*(\gamma,n) \\ U_{(1)}^*(\gamma,k,n) \end{bmatrix}, \quad Q^{(\gamma)}(f) = \begin{bmatrix} Q_{11}^{(\gamma)}(f) & Q_{(1)}^{(\gamma,1)}(f) \\ \mathbf{0} & Q^{(\gamma,1)}(f) \end{bmatrix} \tag{4.11}$$

where $\quad Q_{(1)}^{(\gamma,1)}(f) = \begin{bmatrix} Q_{12}^{(\gamma)}(f) & \dots & Q_{1s}^{(\gamma)}(f) \end{bmatrix},$

$$Q^{(\gamma,1)}(f) = \begin{bmatrix} Q_{22}^{(\gamma)}(f) & \dots & Q_{2s}^{(\gamma)}(f) \\ \vdots & \ddots & \vdots \\ \mathbf{0} & \dots & Q_{ss}^{(\gamma)}(f) \end{bmatrix} \quad \text{and} \quad U_{(1)}^*(\gamma,k,n) = \begin{bmatrix} U_2^*(\gamma,k,n) \\ \vdots \\ U_s^*(\gamma,k,n) \end{bmatrix}.$$

Hence (4.10) can be also written as

$$\begin{bmatrix} U_1^*(\gamma,k,n) \\ U_{(1)}^*(\gamma,k,n) \end{bmatrix} = \max_{f\in\mathcal{F}} \begin{bmatrix} Q_{11}^{(\gamma)}(f) & Q_{(1,1)}^{(\gamma)}(f) \\ \mathbf{0} & Q^{(\gamma,1)}(f) \end{bmatrix} \cdot \begin{bmatrix} U_1^*(\gamma,k+1,n) \\ U_{(1)}^*(\gamma,k+1,n) \end{bmatrix} \tag{4.12}$$

where the structure of the diagonal block $Q_{11}^{(\gamma)}(f)$ is given by (4.2). However, by (4.8)

$$Q_{11}^{(\gamma)}(f)\, x_1^{(\gamma)}(\hat f) \le Q_{11}^{(\gamma)}(\hat f)\, x_1^{(\gamma)}(\hat f) = \rho_1^{(\gamma)}(\hat f)\, x_1^{(\gamma)}(\hat f)$$

with $\rho_1^{(\gamma)}(\hat f) > \rho_2^{(\gamma)}(\hat f)$ (by Assumption 4.1). Let E be the matrix of one's, then for sufficiently small $\varepsilon > 0$ the spectral radius $\rho_{(1)}^{(\gamma,\varepsilon)}(f)$ of an irreducible matrix

$$Q^{(\gamma,1,\varepsilon)}(f) := Q^{(\gamma,1)}(f) + \varepsilon E$$

is less than $\rho_1^{(\gamma)}(f)$; hence $\varepsilon(f) := \rho_{(1)}^{(\gamma,\varepsilon)}(f)/\rho_1^{(\gamma)}(f) < \varepsilon^* < 1$, for any $f \in \mathcal{F}$, and the corresponding right Perron eigenvector $x_{(1)}^{(\gamma,1,\varepsilon)}(f)$ is strictly positive. Then for

$$Q^{(\gamma,\varepsilon)}(f) = \begin{bmatrix} Q_{11}^{(\gamma)}(f) & Q_{(1)}^{(\gamma,1,\varepsilon)}(f) \\ \mathbf{0} & Q^{(\gamma,1,\varepsilon)}(f) \end{bmatrix}$$

we have

$$(\boldsymbol{Q}^{(\gamma,\varepsilon)}(f))^n = \begin{bmatrix} (\boldsymbol{Q}_{11}^{(\gamma)}(f))^n & \sum\limits_{k+\ell=n-1} (\boldsymbol{Q}_{11}^{(\gamma)}(f))^k\, \boldsymbol{Q}_{(1)}^{(\gamma,1,\varepsilon)}(f)\, (\boldsymbol{Q}^{(\gamma,1,\varepsilon)}(f))^\ell \\ \mathbf{0} & (\boldsymbol{Q}^{(\gamma,1,\varepsilon)}(f))^n \end{bmatrix}. \tag{4.13}$$

Moreover, selecting $\boldsymbol{Q}_{(1)}^{(\gamma,1,\varepsilon)}$ such that $\boldsymbol{Q}_{(1)}^{(\gamma,1,\varepsilon)} \geq \boldsymbol{Q}_{(1)}^{(\gamma,1,\varepsilon)}(f)$ for any $f \in \mathcal{F}$ and choosing $\alpha \in \mathbb{R}^+$ such that $\boldsymbol{Q}_{(1)}^{(\gamma,1,\varepsilon)} \cdot \boldsymbol{x}_2^{(\gamma)}(\hat{f}) \leq \alpha\, \boldsymbol{x}_1^{(\gamma)}(\hat{f})$ it holds

$$\begin{aligned} & \sum_{k+\ell=n-1} (\boldsymbol{Q}_{11}^{(\gamma)}(f))^k \cdot \boldsymbol{Q}_{(1,1,\varepsilon)}^{(\gamma)}(f) \cdot (\boldsymbol{Q}^{(\gamma,1,\varepsilon)}(f))^\ell \cdot \boldsymbol{x}_{(1)}^{(\gamma)}(\hat{f}) \\ \leq \;& \alpha \sum_{k+\ell=n-1} (\rho_1^{(\gamma)}(\hat{f}))^k \cdot (\rho^{(\gamma,1)}(\hat{f}))^\ell \boldsymbol{x}_1^{(\gamma)}(\hat{f}) \\ = \;& \alpha\, (\rho_1^{(\gamma)}(\hat{f}))^{n-1} \sum_{\ell=0}^{n-2} (\varepsilon^*)^\ell \boldsymbol{x}_1^{(\gamma)}(\hat{f}) \leq \alpha\, (\rho_{(1)}^{(\gamma)}(\hat{f}))^{n-1} \cdot \frac{1}{1-\varepsilon^*} \cdot \boldsymbol{x}_1^{(\gamma)}(\hat{f}). \end{aligned}$$

Observe that the above bounds hold also for nonhomogeneous products of matrices $\boldsymbol{Q}^{(\gamma,\varepsilon)}(f)$ if $f \in \mathcal{F}$.
Hence if $\boldsymbol{x}_1^{(\gamma)}(f) \geq \boldsymbol{e}$, $\boldsymbol{x}_{(1)}^{(\gamma)}(f) \geq \boldsymbol{e}$ we have for policy π^* fulfilling the dynamic programming recursion (4.12)

$$\begin{bmatrix} \boldsymbol{U}_1^*(\gamma,k,n) \\ \boldsymbol{U}_{(1)}^*(\gamma,k,n) \end{bmatrix} \leq \begin{bmatrix} (\rho_1^{(\gamma)}(\hat{f}))^{n-k} \left\{ \rho_1^{(\gamma)}(\hat{f}) + \alpha \frac{1}{1-\varepsilon^*} \cdot \right\} \cdot \boldsymbol{x}_1^{(\gamma)}(\hat{f}) \\ (\rho^{(\gamma,1)}(\hat{f}))^{n-k} \cdot \boldsymbol{x}_{(1)}^{(\gamma)}(\hat{f}) \end{bmatrix}$$

and the maximal possible growth of $\boldsymbol{U}_1^*(\gamma,n)$ is dominated by $\rho_1^{(\gamma)}(\hat{f})$.

So we have arrived at the following

Theorem 4.2. Let Assumption 4.1 hold. Then for the matrix $\boldsymbol{Q}^{(\gamma)}(\hat{f})$ with $\hat{f} \in \mathcal{F}$ decomposed in accordance with the basic partition $\hat{\mathcal{I}}$ of the state space it holds:
Maximum possible growth rate $G_j^{\hat{\pi}}(\gamma)$ is the same for each $j \in \mathcal{I}_i(\hat{f})$ and is equal to $\rho_i^{(\gamma)}(\hat{f})$. Moreover, this growth rate can be obtained if stationary policy $\pi \sim (\hat{f})$ is followed.

Since by (1.4), (1.7) and (1.8) for each $j \in \mathcal{I}$

$$J_j^{\hat{\pi}}(\gamma) = \frac{1}{\gamma} \lim_{n\to\infty} \frac{1}{n}\, U_j^{\hat{\pi}}(\gamma,n) \tag{4.14}$$

if $j \in \mathcal{I}_i(\hat{f})$ we have for suitably selected $\boldsymbol{x}_i^{(\gamma)}(\hat{f})$ with elements $x_j^{(\gamma)}(\hat{f})$

$$J_j^{\hat{\pi}}(\gamma) = \frac{1}{\gamma} \lim_{n\to\infty} \frac{1}{n} \ln[(\rho_i(\hat{f}))^n \cdot x_j(\hat{f}))] = \frac{1}{\gamma} \ln[\rho_i(\hat{f})]. \tag{4.15}$$

So by Theorem 4.2 for maximal average optimality in risk sensitive Markov decision processes we have the following result.

Theorem 4.3. Let Assumption 4.1 hold. Considering the basic partition of the state space $\hat{\mathcal{I}} = \mathcal{I}_1(\hat{f}) \cup \mathcal{I}_2(\hat{f}) \cup \ldots \cup \mathcal{I}_s(\hat{f})$ it holds:
Maximum average rewards $J_j^{\hat{\pi}}(\gamma)$ are the same for each $j \in \mathcal{I}_i(\hat{f})$ and are equal to $(\gamma)^{-1} \ln[\rho_i^{(\gamma)}(\hat{f})]$.

Example 2. Consider an uncontrolled model (hence the argument f is deleted) where $N = 6$ with transition probability matrix $\boldsymbol{P}$ and transition reward matrix $\boldsymbol{R}$ (with elements r_{ij}'s) given by:

$$\boldsymbol{P} = \begin{bmatrix} 0.25 & 0.25 & 0.25 & 0 & 0.25 & 0 \\ 0.25 & 0.25 & 0 & 0.25 & 0 & 0.25 \\ 0 & 0 & 0.1 & 0.1 & 0.4 & 0.4 \\ 0 & 0 & 0.1 & 0.1 & 0.4 & 0.4 \\ 0 & 0 & 0 & 0 & 0.5 & 0.5 \\ 0 & 0 & 0 & 0 & 0.5 & 0.5 \end{bmatrix}, \quad \boldsymbol{R} = \begin{bmatrix} 1 & 1 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 2 & 0 & 0 \\ 0 & 0 & 2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}.$$

Obviously, the transition probability matrix $\boldsymbol{P}$ contains a single class $\boldsymbol{P}_{(\mathrm{BB})}$ of recurrent states and two irreducible classes of transient states $\boldsymbol{P}_{(\mathrm{N})11}$ and $\boldsymbol{P}_{(\mathrm{N})22}$ where

$$\boldsymbol{P}_{(\mathrm{BB})} = \begin{bmatrix} 0.5 & 0.5 \\ 0.5 & 0.5 \end{bmatrix}, \quad \boldsymbol{P}_{(\mathrm{N})11} = \begin{bmatrix} 0.25 & 0.25 \\ 0.25 & 0.25 \end{bmatrix}, \quad \boldsymbol{P}_{(\mathrm{N})22} = \begin{bmatrix} 0.1 & 0.1 \\ 0.1 & 0.1 \end{bmatrix}$$

and hence

$$\boldsymbol{Q}^{(\gamma)} = \begin{bmatrix} \boldsymbol{Q}^{(\gamma)}_{(\mathrm{N})11} & \boldsymbol{Q}^{(\gamma)}_{(\mathrm{N})12} & \boldsymbol{Q}^{(\gamma)}_{(\mathrm{N})1(\mathrm{B})} \\ \boldsymbol{0} & \boldsymbol{Q}^{(\gamma)}_{(\mathrm{N})22} & \boldsymbol{Q}^{(\gamma)}_{(\mathrm{N})2(\mathrm{B})} \\ \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{Q}^{(\gamma)}_{(\mathrm{BB})} \end{bmatrix} \quad \text{with} \quad \boldsymbol{Q}^{(\gamma)}_{(\mathrm{BB})} = \begin{bmatrix} 0.5 & 0.5 \\ 0.5 & 0.5 \end{bmatrix},$$

and the spectral radius $\rho^{(\gamma)}_{(\mathrm{B})}$ of $\boldsymbol{Q}^{(\gamma)}_{(\mathrm{BB})}$ is equal to one for any γ.

However, for the remaining two diagonal submatrices of $\boldsymbol{Q}^{(\gamma)}$ we have

$$\boldsymbol{Q}^{(\gamma)}_{(\mathrm{N})11} = \begin{bmatrix} 0.25\,\mathrm{e}^{\gamma} & 0.25\,\mathrm{e}^{\gamma} \\ 0.25\,\mathrm{e}^{\gamma} & 0.25\,\mathrm{e}^{\gamma} \end{bmatrix}, \qquad \boldsymbol{Q}^{(\gamma)}_{(\mathrm{N})22} = \begin{bmatrix} 0.1 & 0.1\,\mathrm{e}^{2\gamma} \\ 0.1\,\mathrm{e}^{2\gamma} & 0.1 \end{bmatrix}.$$

After some algebra we conclude that the spectral radius $\rho^{(\gamma)}_{(\mathrm{N})1}$ of $\boldsymbol{Q}^{(\gamma)}_{(\mathrm{N})11}$ is equal to $0.5\,\mathrm{e}^{\gamma}$ and the spectral radius $\rho^{(\gamma)}_{(\mathrm{N})2}$ of $\boldsymbol{Q}^{(\gamma)}_{(\mathrm{N})22}$ is equal to $0.1\,(\mathrm{e}^{2\gamma}+1)$. Obviously, for every $\boldsymbol{Q}^{(\gamma)}_{(\mathrm{BB})}$, $\boldsymbol{Q}^{(\gamma)}_{(\mathrm{N})11}$, $\boldsymbol{Q}^{(\gamma)}_{(\mathrm{N})22}$ the corresponding (right) Perron eigenvector is the two-dimensional unit vector.
Obviously, both $0.5\,\mathrm{e}^{\gamma}$ and $0.1\,(\mathrm{e}^{2\gamma}+1)$ are increasing functions of the risk aversion coefficient γ; moreover $0.5\,\mathrm{e}^{\gamma} = 1$ for $\gamma = \ln 2$, $0.1\,(\mathrm{e}^{2\gamma}+1) = 1$ for $\gamma = \frac{1}{2}\ln 9$, and $0.1\,(\mathrm{e}^{2\gamma}+1) = 0.5\,\mathrm{e}^{\gamma}$ for $\gamma = \ln\left(\frac{5+\sqrt{21}}{2}\right)$.

So:

For $\gamma < \ln 2$ it holds $\rho^{(\gamma)}_{(\mathrm{B})} > \rho^{(\gamma)}_{(\mathrm{N})1} > \rho^{(\gamma)}_{(\mathrm{N})2}$, $\rho^{(\gamma)} = \rho^{(\gamma)}_{(\mathrm{B})} = 1$ and the right Perron eigenvector $\boldsymbol{x}^{(\gamma)}$ of $\boldsymbol{Q}^{(\gamma)}$ is strictly positive.

For $\gamma \in \big(\ln 2, \frac{1}{2}\ln 9)\big)$ it holds $\rho^{(\gamma)}_{(\mathrm{N})1} > \rho^{(\gamma)}_{(\mathrm{B})} > \rho^{(\gamma)}_{(\mathrm{N})2}$, $\rho^{(\gamma)} = \rho^{(\gamma)}_{(\mathrm{N})1} = 0.5\mathrm{e}^{\gamma}$ and the right Perron eigenvector $\boldsymbol{x}^{(\gamma)}$ of $\boldsymbol{Q}^{(\gamma)}$ is not strictly positive.

For $\gamma \in \left(\frac{1}{2}\ln 9, \ln\left(\frac{5+\sqrt{21}}{2}\right)\right)$ it holds $\rho^{(\gamma)}_{(\mathrm{N})1} > \rho^{(\gamma)}_{(\mathrm{N})2} > \rho^{(\gamma)}_{(\mathrm{B})}$, $\rho^{(\gamma)} = \rho^{(\gamma)}_{(\mathrm{N})1} = 0.5\mathrm{e}^{\gamma}$ and the right Perron eigenvector $\boldsymbol{x}^{(\gamma)}$ of $\boldsymbol{Q}^{(\gamma)}$ is not strictly positive.

For $\gamma > \ln(\frac{5+\sqrt{21}}{2})$ it holds $\rho^{(\gamma)}_{(\mathrm{N})2} > \rho^{(\gamma)}_{(\mathrm{N})1} > \rho^{(\gamma)}_{(\mathrm{B})}$, $\rho^{(\gamma)} = \rho^{(\gamma)}_{(\mathrm{N})2} = 0.1(\mathrm{e}^{2\gamma}+1)$ and the right Perron eigenvector $\boldsymbol{x}^{(\gamma)}$ of $\boldsymbol{Q}^{(\gamma)}$ is not strictly positive.

(The Perron eigenvectors are explicitly calculated – see the footnote.)

Hence[2] for the growth rate and average rewards we have:

If $\gamma < \ln 2$ then the growth rate $G_i(\gamma) = 1$ and average reward $J_i(\gamma) = 0$ for an arbitrary starting state i.

For $\gamma \in \big(\ln 2, \frac{1}{2}\ln 9\big)$ the growth rate $G_i(\gamma) = 1$ and average reward $J_i(\gamma) = 0$ for starting states $i = 3,4,5,6$. If the chain starts in state $i = 1,2$ then the growth rate $G_i(\gamma) = 0.5\,\mathrm{e}^{\gamma}$ and average reward $J_i(\gamma) = 1 - \frac{1}{\gamma}\ln 2$.

For $\gamma \in \left(\frac{1}{2}\ln 9, \ln\left(\frac{5+\sqrt{21}}{2}\right)\right)$ the growth rate $G_i(\gamma) = 1$ and average reward $J_i(\gamma) = 0$ only for starting states $i = 5,6$. If the chain starts in states $i = 3,4$ then the growth rate $G_i(\gamma) = 0.1\,(\mathrm{e}^{2\gamma}+1)$ and average reward $J_i(\gamma) = \frac{1}{\gamma}\ln[0.1\,(\mathrm{e}^{2\gamma}+1)]$. If the chain starts in state $i = 1,2$ then the growth rate $G_i(\gamma) = 0.5\,\mathrm{e}^{\gamma}$ and average reward $J_i(\gamma) = 1 - \frac{1}{\gamma}\ln 2$.

If $\gamma > \ln\left(\frac{5+\sqrt{21}}{2}\right)$ the growth rate $G_i(\gamma) = 1$ and average reward $J_i(\gamma) = 0$ only for starting states $i = 5,6$. However, if the chain starts in state $i = 1,2,3,4$ then the growth rate $G_i(\gamma) = 0.1\,(\mathrm{e}^{2\gamma}+1)$ and average reward $J_i(\gamma) = \frac{1}{\gamma}\ln[0.1\,(\mathrm{e}^{2\gamma}+1)]$.

5. SIMPLE MODELS WITH MINIMAL COSTS

In this section we show that the results of Section 3 can be easily extended to risk-sensitive Markov decision chains where instead of transition rates we consider transition costs $c_{ij}(a)$ and our aim is to minimizes the growth rates and the corresponding average costs.

In parallel to Assumption 3.1 we make

---

[2] After some algebra we obtain for the right Perron eigenvectors

If $\gamma < \ln 2$ then $\boldsymbol{x}^{(\gamma)} = \left[0.25\frac{17-\mathrm{e}^{2\gamma}}{(1-0.5\mathrm{e}^{\gamma})(9-\mathrm{e}^{2\gamma})} \;\; 0.25\frac{17-\mathrm{e}^{2\gamma}}{(1-0.5\mathrm{e}^{\gamma})(9-\mathrm{e}^{2\gamma})} \;\; \frac{8}{9-\mathrm{e}^{2\gamma}} \;\; \frac{8}{9-\mathrm{e}^{2\gamma}} \;\; 1 \;\; 1\right]^{\mathrm{T}}$;

If $\gamma \in \big(\ln 2, \frac{1}{2}\ln 9\big)$ then $\boldsymbol{x}^{(\gamma)} = [1\;\;1\;\;0\;\;0\;\;0\;\;0]^{\mathrm{T}}$;

If $\gamma \in \left(\frac{1}{2}\ln 9, \ln(\frac{5+\sqrt{21}}{2})\right)$ then $\boldsymbol{x}^{(\gamma)} = [1\;\;1\;\;0\;\;0\;\;0\;\;0]^{\mathrm{T}}$;

If $\gamma > \ln\left(\frac{5+\sqrt{21}}{2}\right)$ then $\boldsymbol{x}^{(\gamma)} = \left[\frac{1}{0.4(\mathrm{e}^{2\gamma}+1)-2\mathrm{e}^{\gamma}} \;\; \frac{1}{0.4(\mathrm{e}^{2\gamma}+1)-2\mathrm{e}^{\gamma}} \;\; 1 \;\; 1 \;\; 0 \;\; 0\right]^{\mathrm{T}}$.

Assumption 5.1. There exists $\hat{\rho}^{(\gamma)} \equiv \rho^{(\gamma)}(\hat{f})$ and $\widehat{\boldsymbol{x}}^{(\gamma)} \equiv \boldsymbol{x}^{(\gamma)}(\hat{f}) > \boldsymbol{0}$ (unique up to a multiplicative constant) such that (for a given value of the risk aversion coefficient γ)

$$\hat{\rho}^{(\gamma)}\, \widehat{\boldsymbol{x}}^{(\gamma)} = \min_{f\in\mathcal{F}}\{\boldsymbol{Q}^{(\gamma)}(f)\cdot\widehat{\boldsymbol{x}}^{(\gamma)}\} = \boldsymbol{Q}^{(\gamma)}(\hat{f})\cdot\widehat{\boldsymbol{x}}^{(\gamma)} \tag{5.1}$$

such that on using the matrix decomposition according to (2.9), i. e., on writing

$$\boldsymbol{Q}^{(\gamma)}(\hat{f}) = \begin{bmatrix} \boldsymbol{Q}^{(\gamma)}_{(\mathrm{NN})}(\hat{f}) & \boldsymbol{Q}^{(\gamma)}_{(\mathrm{NB})}(\hat{f}) \\ \boldsymbol{0} & \boldsymbol{Q}^{(\gamma)}_{(\mathrm{BB})}(\hat{f}) \end{bmatrix} \tag{5.2}$$

where $\boldsymbol{Q}^{(\gamma)}_{(\mathrm{BB})}(\hat{f})$ is the "biggest" diagonal class with spectral radius $\rho^{(\gamma)}(\hat{f})$ among all $\boldsymbol{Q}^{(\gamma)}(f)$'s fulfilling (5.1).

Remark. Of course, a sufficient condition for existence of the matrix $\boldsymbol{Q}^{(\gamma)}(\hat{f})$ fulfilling condition (5.1) is the existence of $\boldsymbol{x}^{(\gamma)}(f) > \boldsymbol{0}$ for each $f \in \mathcal{F}$. As it is shown in the Appendix using e. g. policy iterations in a finite number of steps we can find $\hat{f} \in \mathcal{F}$ such that $\rho^{(\gamma)}(\hat{f}) \equiv \hat{\rho}^{(\gamma)}$ is the minimal possible eigenvalue of $\boldsymbol{Q}^{(\gamma)}(f)$ over all $f \in \mathcal{F}$, and

$$\begin{aligned} \boldsymbol{Q}^{(\gamma)}(f)\cdot\boldsymbol{x}^{(\gamma)}(\hat{f}) \;&\geq\; \min_{f\in\mathcal{F}}\{\boldsymbol{Q}^{(\gamma)}(f)\cdot\boldsymbol{x}^{(\gamma)}(\hat{f})\} \\ = \boldsymbol{Q}^{(\gamma)}(\hat{f})\cdot\boldsymbol{x}^{(\gamma)}(\hat{f}) \;&=\; \rho^{(\gamma)}(\hat{f})\,\boldsymbol{x}^{(\gamma)}(\hat{f}), \quad \text{with} \;\; \boldsymbol{x}^{(\gamma)}(\hat{f}) > \boldsymbol{0}. \end{aligned} \tag{5.3}$$

Theorem 5.1. If condition (5.1) holds then for a given value of the risk aversion coefficient γ there exist numbers $\beta_1^{(\gamma)} > \beta_2^{(\gamma)} > 0$ such that

$$\beta_1^{(\gamma)}\,\boldsymbol{x}^{(\gamma)}(\hat{f}) \leq (\hat{\rho}^{(\gamma)})^{-n}\prod_{k=0}^{n-1}\boldsymbol{Q}^{(\gamma)}(f^k)\cdot\boldsymbol{e} \leq \beta_2^{(\gamma)}\,\boldsymbol{x}^{(\gamma)}(\hat{f}) \tag{5.4}$$

for any policy $\pi = (f^k)$ minimizing the growth of $\boldsymbol{U}^{\pi}(\gamma, n)$ for $n = 0, 1, \ldots$. In addition, (5.4) is also fulfilled for stationary policy $\hat{\pi} \sim (\hat{f})$.

Theorem 5.1 can be rephrased in words as

Corollary 5.2. Under condition (5.1) if policy $\pi = (f^n)$ minimizing $\boldsymbol{U}^{\pi}(\gamma, n)$ is followed, the *growth rate* of each element of $\boldsymbol{U}^{\pi}(\gamma, n) = \prod_{k=0}^{n-1}\boldsymbol{Q}^{(\gamma)}(f^k)\cdot\boldsymbol{e}$ is the same and equals $\hat{\rho}^{(\gamma)}$. Moreover, stationary policy $\hat{\pi} \sim (\hat{f})$ also minimizes the growth rate.

Denoting elements of $\boldsymbol{x}^{(\gamma)}(\hat{f}) > \boldsymbol{0}$ by $x_j^{(\gamma)}(f)$ (for $j = 1, \ldots, N$) and elements of an $N\times N$ matrix $\boldsymbol{Q}^{(\gamma)}(f)$ by $q_{ij}^{(\gamma)}(f_i)$ (recall that by (1.10) $q_{ij}^{(\gamma)}(\cdot) = p_{ij}(\cdot)\cdot \mathrm{e}^{\gamma r_{ij}(\cdot)}$), (5.1) can be also written in alternative forms for $g(f) := \gamma^{-1} \ln(\rho^{(\gamma)}(f))$ and $w_j(f) := \gamma^{-1} \ln(x_j^{(\gamma)}(f))$ (for $j = 1, \ldots, N$) as the following set of (nonlinear) equations:

$$\mathrm{e}^{\gamma\,(g(\hat{f})+w_i(\hat{f}))} = \min_{a\in\mathcal{A}_i} \left\{ \sum_{j\in\mathcal{I}} p_{ij}(a) \cdot \mathrm{e}^{\gamma\,(r_{ij}(a)+w_j(\hat{f}))} \right\}, \quad \text{for} \quad i = 1, \ldots, N, \tag{5.5}$$

called γ-*average cost optimality equation*.
In the multiplicative form (5.5) takes on the form:

$$\rho^{(\gamma)}(\hat{f})\, x_i^{(\gamma)}(\hat{f}) = \min_{a\in\mathcal{A}_i} \left\{ \sum_{j\in\mathcal{I}} p_{ij}(a) \cdot \mathrm{e}^{\gamma r_{ij}} \cdot x_j^{(\gamma)}(\hat{f}) \right\}, \quad \text{for} \quad i = 1, \ldots, N. \tag{5.6}$$

Observe that the solution to (5.5), resp. (5.6), i. e. $g(\hat{f})$, $w_i(\hat{f})$, resp. $\rho^{(\gamma)}(\hat{f})$, $x_i^{(\gamma)}(\hat{f})$ is unique up to an additive constant (added to $w_i(\hat{f})$'s), resp. multiplicative constant (applied to $x_i^{(\gamma)}(\hat{f})$'s).

Using the above facts Corollary 5.2 can be formulated as

Theorem 5.3. If condition (5.1) holds then for any policy $\pi = (f^n)$ the asymptotical mean value $J_i^\pi(\gamma, 0)$ is bounded from below by $g(\hat{f}) := \gamma^{-1} \ln(\rho^{(\gamma)}(\hat{f}))$. Moreover, stationary policy $\hat{\pi} \sim (\hat{f})$ yields the minimal asymptotical mean value $J_i^\pi(\gamma, 0)$ that is independent of the starting state $i \in \mathcal{I}$ and equal to $g(\hat{f}) := \gamma^{-1} \ln(\rho^{(\gamma)}(\hat{f}))$.

6. CONCLUSIONS

The paper, inspired by the work of R. Cavazos-Cadena and D. Hernández-Hernández [9, 10], presents a complete characterization of policies maximizing growth rates and the mean values of the associated certainty equivalents over an infinite time horizon in risk-sensitive Markov decision chains with finite state and action spaces.

The study of the type of dynamic programming problem was initiated by Bellman in [1, 2]. For an "easy case" when the underlying Markov chain contains a single class of recurrent state and no transient states (i. e. all states are communicating) and the transition probability matrix $\boldsymbol{P}(f)$ is irreducible for any $f \in \mathcal{F}\}$ also the nonnegative matrices $\boldsymbol{Q}^{(\gamma)}(f)$'s obtained from the transition probability matrices $\boldsymbol{P}(f)$'s must be irreducible. Then for each $\boldsymbol{Q}^{(\gamma)}(f)$ with $f \in \mathcal{F}$ there exists strictly positive (right or left) Perron eigenvector and using policy iterations we can find $\{\hat{f} \in \mathcal{F}\}$ such that (5.1) or (5.3) holds. Then for each $f \in \mathcal{F}$ the growth rate of $\boldsymbol{Q}^{(\gamma)}(f)$ and also the corresponding values of certainty equivalents are independent of the starting state (cf. [12, 16]).

Moreover, these results can also be extended to models with transient states. As we have shown in case that there exists strictly positive right Perron eigenvector the growth rate of $\boldsymbol{Q}^{(\gamma)}(f)$ and also the corresponding values of certainty equivalents are still independent of the starting state. This holds both maximal or minimal growth rates and maximal or minimal values of certainty equivalents.

Unfortunately, in the general case no strictly positive right Perron eigenvector need not exist. In this case, if we maximize the growth rates and also the corresponding values of certainty equivalents the block-triangular decomposition of the collection of nonnegative matrices $\{\boldsymbol{Q}^{\gamma}(f), f \in \mathcal{F}\}$ may be very helpful. For details see [15, 17, 18, 19, 20] where algorithmic procedures for finding block-triangular decomposition of the matrix set $\{\boldsymbol{Q}^{(\gamma)}(f), f \in \mathcal{F}\}$ fulfilling conditions (4.6) – (4.8) were suggested. This approach is a bit technical, but as we have shown, it enables to identify subsets of starting states with the maximal growth rate of $\boldsymbol{Q}^{\gamma}(f)$ and also maximal values of certainty equivalents. In addition, in an early paper Mandl [14] investigates convergence radius (i. e. the reciprocal value of the growth rate) of nonhomogeneous products of a collection of (reducible) nonnegative matrices arising in dynamic programming using different methods of ours.

Problems of this type began again very popular in the last ten years (see e. g. [4] – [10]), however, there were not intensively studied in connection with nonnegative matrices.

A companion problem of finding policies that minimize the growth rates and also the corresponding values of certainty equivalents for the risk-sensitive Markov control processes with reducible transition probability matrices is under current research.

APPENDIX: ON A POLICY ITERATION METHOD

For the sake of completeness we present policy iteration algorithm for finding a matrix with minimal possible eigenvalue in the class of nonnegative matrices with strictly positive (right) Perron eigenvectors along with its concise proof. For the irreducible case (guaranteeing existence of a strictly positive Perron eigenvectors) the algorithm along with its proof is strictly similar to the procedure suggested in Howard and Matheson [12] for finding a matrix with maximal positive eigenvalue (it suffices to change min to max). However, some extensions are necessary for handling the case with reducible matrices possessing strictly positive right eigenvectors. For the sake of simplicity we shall omit the superscript (γ).

Algorithm A.

Step 0. Select matrix $\boldsymbol{Q}(f^{(0)})$ with $f^{(0)} \in \mathcal{F}$ such that the row sums are minimal, i. e., it holds $\boldsymbol{Q}(f^{(0)}) \cdot \boldsymbol{e} \le \boldsymbol{Q}(f) \cdot \boldsymbol{e}$ for any $f \in \mathcal{F}$.

Step 1. For the matrix $\boldsymbol{Q}(f^{(k)})$ with $f^{(k)} \in \mathcal{F}$, $k = 0, 1, \ldots$ calculate its spectral radius $\rho(f^{(k)})$ along with its right Perron eigenvector $\boldsymbol{x}(f^{(k)})$, cf. (2.8).

Step 2. Construct (if possible) the matrix $\boldsymbol{Q}(f^{(k+1)})$ with $f^{(k+1)} \in \mathcal{F}$, such that

$$\boldsymbol{Q}(f^{(k+1)}) \cdot \boldsymbol{x}(f^{(k)}) < \rho(f^{(k)})\, \boldsymbol{x}(f^{(k)}) = \boldsymbol{Q}(f^{(k)}) \cdot \boldsymbol{x}(f^{(k)}) \qquad \text{(A.1)}$$

(i. e., a strict inequality holds at least for one $i \in \mathcal{I}$).

Step 3. If such a matrix $\boldsymbol{Q}(f^{(k+1)})$ exists, then set $\boldsymbol{Q}(f^{(k+1)}) := \boldsymbol{Q}(f^{(k)})$ and repeat Step 1, else set $\widehat{\boldsymbol{Q}} := \boldsymbol{Q}(f^{(k)})$, $\hat{f} := f^{(k)}$ and stop.

Theorem A. The sequence of spectral radii $\rho(f^{(k)})$ generated by Algorithm A is non-increasing (i. e. $\rho(f^{(k+1)} \le \rho(f^{(k)}))$, resp. decreasing if $\boldsymbol{Q}(f^{(k+1)})$ is irreducible, and the sequence $\boldsymbol{Q}(f^{(k)})$ converges monotonously to the matrix $\widehat{\boldsymbol{Q}} = \boldsymbol{Q}(\hat{f})$ such that

$$\boldsymbol{Q}(f) \cdot \boldsymbol{x}(\hat{f}) \quad \ge \quad \rho(\hat{f})\, \boldsymbol{x}(\hat{f}) = \boldsymbol{Q}(\hat{f}) \cdot \boldsymbol{x}(\hat{f}), \quad \text{with} \quad \boldsymbol{x}(\hat{f}) > \mathbf{0} \tag{A.2}$$

$$\rho(f) \quad \ge \quad \rho(\hat{f}) \equiv \hat{\rho} \qquad \text{for all} \quad f \in \mathcal{F}. \tag{A.3}$$

P r o o f. Employing policy iterations in accordance with Algorithm A we are able to show that

$$\rho(f^{(k+1)}) \le \rho(f^{(k)}) \quad \text{for} \quad k = 0, 1, \dots \tag{A.4}$$

$$\rho(f^{(k+1)}) = \rho(f^{(k)}) \Rightarrow \boldsymbol{x}(f^{(k+1)}) \le \boldsymbol{x}(f^{(k)}) \quad \text{with} \tag{A.5}$$

$$x_i(f^{(k+1)}) \;=\; x_i(f^{(k)}) \;\; \text{for all } i \in \mathcal{I} \;\; \text{pertaining to any basic class of} \;\; \boldsymbol{Q}(f^{(k+1)}).$$

To this end observe that by (2.8)

$$\rho(f^{(k+1)})\, \boldsymbol{x}(f^{(k+1)}) - \rho(f^{(k)})\, \boldsymbol{x}(f^{(k)}) = \boldsymbol{Q}(f^{(k+1)}) \cdot \boldsymbol{x}(f^{(k+1)}) - \boldsymbol{Q}(f^{(k)}) \cdot \boldsymbol{x}(f^{(k)})$$

and after some algebra we conclude that

$$\begin{aligned} &\Big[\rho(f^{(k+1)}) - \rho(f^{(k)})\Big] \cdot \boldsymbol{x}(f^{(k)}) + \rho(f^{(k+1)}) \, \Big[\boldsymbol{x}(f^{(k+1)}) - \boldsymbol{x}(f^{(k)})\Big] \\ &= \boldsymbol{Q}(f^{(k+1)}) \cdot \Big[\boldsymbol{x}(f^{(k+1)}) - \boldsymbol{x}(f^{(k)})\Big] + \Big[\boldsymbol{Q}(f^{(k+1)}) - \boldsymbol{Q}(f^{(k)})\Big] \cdot \boldsymbol{x}(f^{(k)}). \end{aligned} \tag{A.6}$$

On premultiplying (A.6) by the left Perron eigenvector $\boldsymbol{y}(f^{(k+1)}) \ge \mathbf{0}$ and recalling that by (A.1)

$$\boldsymbol{\varphi}(f^{(k+1)}, f^{(k)}) := \Big[\boldsymbol{Q}(f^{(k+1)}) - \boldsymbol{Q}(f^{(k)})\Big] \cdot \boldsymbol{x}(f^{(k)}) < \mathbf{0} \tag{A.7}$$

we immediately conclude that by (2.7) also

$$[\rho(f^{(k+1)}) - \rho(f^{(k)})]\, \boldsymbol{y}(f^{(k+1)}) \cdot \boldsymbol{x}(f^{(k+1)}) = \boldsymbol{y}(f^{(k+1)}) \cdot \boldsymbol{\varphi}(f^{(k+1)}, f^{(k)}) \le \mathbf{0} \tag{A.8}$$

implying $\rho(f^{(k+1)}) \le \rho(f^{(k)})$ with $\rho(f^{(k+1)}) = \rho(f^{(k)})$ iff $\varphi_i(f^{(k+1)}, f^{(k)}) = 0$ for all $i \in \mathcal{I}$ pertaining to any basic class of $\boldsymbol{Q}(f^{(k+1)})$.

In particular, for $\boldsymbol{z}(f^{(k+1)}) := \boldsymbol{x}(f^{(k+1)}) - \boldsymbol{x}(f^{(k)})$ in case that $\rho(f^{(k+1)}) = \rho(f^{(k)})$ Eq. (A.6) can also written as

$$\rho(f^{(k)})\boldsymbol{z}(f^{(k+1)}) = \boldsymbol{Q}(f^{(k+1)}) \cdot \boldsymbol{z}(f^{(k+1)}) + \boldsymbol{\varphi}(f^{(k+1)}, f^{(k)}). \tag{A.9}$$

If we decompose $\boldsymbol{Q}(f^{(k+1)})$ according to (2.9) and apply this decomposition to (A.9)we easily verify that $\boldsymbol{x}_{(\mathrm{B})}(f^{(k+1)}) = \boldsymbol{x}_{(\mathrm{B})}(f^{(k)})$ and $\boldsymbol{\varphi}_{(\mathrm{B})}(f^{(k+1)}, f^{(k)}) = \mathbf{0}$. Then by (A.9) we have

$$\begin{aligned} \rho(f^{(k+1)})\, \boldsymbol{z}_{(\mathrm{N})}(f^{(k+1)}) &= \boldsymbol{Q}_{(\mathrm{NN})}(f^{(k+1)}) \cdot \boldsymbol{z}_{(\mathrm{N})}(f^{(k+1)}) + \boldsymbol{\varphi}_{(\mathrm{N})}(f^{(k+1)}, f^{(k)}) \\ \Longrightarrow \quad \boldsymbol{z}_{(\mathrm{N})}(f^{(k+1)}) &= (\rho(f^{(k+1)}))^{-1}[\boldsymbol{I} - (\rho(f^{(k+1)}))^{-1} \cdot \boldsymbol{Q}_{(\mathrm{NN})}(f^{(k+1)})]^{-1} \cdot \\ &\qquad \boldsymbol{\varphi}_{(\mathrm{N})}(f^{(k+1)}, f^{(k)}) < \mathbf{0} \end{aligned} \tag{A.10}$$

and $\boldsymbol{x}(f^{(n+1)}) < \boldsymbol{x}(f^{(n)})$; hence the algorithm cannot cycle.

Since the set $\mathcal{F}$ of decision vectors is finite, Algorithm A terminates in a finite number of steps. □

ACKNOWLEDGEMENT

This work was supported by the Czech Science Foundation under Grants 402/08/0107 and 402/07/1113. The author is indebted to anonymous referee for insightful reading the manuscript and many helpful comments.

(Received October 30, 2007.)

REFERENCES

[1] R. Bellman: On a quasi-linear equation. Canadian J. Math. *8* (1956), 198–202.

[2] R. Bellman: Dynamic Programming. Princenton Univ. Press, Princenton, N.J. 1957.

[3] A. Berman and R. J. Plemmons: Nonnegative Matrices in the Mathematical Sciences. Academic Press, New York 1979.

[4] T. D. Bielecki, D. Hernández-Hernández, and S. R. Pliska: Risk-sensitive control of finite state Markov chains in discrete time, with application to portfolio management. Math. Methods Oper. Res. *50* (1999), 167–188.

[5] R. Cavazos-Cadena: Solution to the risk-sensitive average cost optimality equation in a class of Markov decision processes with finite state space. Math. Methods Oper. Res. *57* (2003), 253–285.

[6] R. Cavazos-Cadena and R. Montes-de-Oca: The value iteration algorithm in risk-sensitive average Markov decision chains with finite state space. Math. Oper. Res. *28* (2003), 752–756.

[7] R. Cavazos-Cadena and R. Montes-de-Oca: Nonstationary value iteration in controlled Markov chains with risk-sensitive average criterion. J. Appl. Probab. *42* (2005), 905–918.

[8] R. Cavazos-Cadena and D. Hernández-Hernández: Solution fo the risk-sensitive average optimality equation in communicating Markov decision chains with finite state space: An alternative approach. Math. Methods Oper. Res. *56* (2002), 473–479.

[9] R. Cavazos-Cadena and D. Hernández-Hernández: A characterization exponential functionals in finite Markov chains. Math. Methods Oper. Res. *60* (2004), 399–414.

[10] R. Cavazos-Cadena and D. Hernández-Hernández: A characterization of the optimal risk-sensitive average cost in finite controlled Markov chains. Ann. Appl. Probab. *15* (2005), 175–212.

[11] F. R. Gantmakher: The Theory of Matrices. Chelsea, London 1959.

[12] R. A. Howard and J. Matheson: Risk-sensitive Markov decision processes. Manag. Sci. *23* (1972), 356–369.
[13] S. A. Jaquette: A utility criterion for Markov decision processes. Manag. Sci. *23* (1976), 43–49.
[14] P. Mandl: Decomposable non-negative matrices in a dynamic programming problem. Czechoslovak Math. J. *20* (1970), 504–510.
[15] U. G. Rothblum and P. Whittle: Growth optimality for branching Markov decision chains. Math. Oper. Res. *7* (1982), 582–601.
[16] K. Sladký: On dynamic programming recursions for multiplicative Markov decision chains. Math. Programming Study *6* (1976), 216–226.
[17] K. Sladký: Successive approximation methods for dynamic programming models. In: Proc. of the Third Formator Symposium on the Analysis of Large-Scale Systems (J. Beneš and L. Bakule, eds.). Academia, Prague 1979, pp. 171–189.
[18] K. Sladký: Bounds on discrete dynamic programming recursions I. Kybernetika *16* (1980), 526–547.
[19] P. Whittle: Optimization Over Time – Dynamic Programming and Stochastic Control. Volume II, Chapter 35, Wiley, Chichester 1983.
[20] W. H. M. Zijm: Nonnegative Matrices in Dynamic Programming. Mathematical Centre Tract, Amsterdam 1983.

Karel Sladký, Institute of Information Theory and Automation – Academy of Sciences of the Czech Republic, Pod Vodárenskou věží 4, 182 08 Praha 8. Czech Republic.
e-mail: sladky@utia.cas.cz