CENTRAL MOMENTS AND RISK-SENSITIVE OPTIMALITY IN MARKOV REWARD CHAINS

Karel Sladký
Institute of Information Theory and Automation of the AS CR

Abstract There is no doubt that usual optimization criteria examined in the literature on optimization of Markov reward processes, e.g. total discounted or mean reward, may be quite insufficient to characterize the problem from the point of the decision maker. To this end it is necessary to select more sophisticated criteria that reflect also the variability-risk features of the problem (cf. Cavazos-Cadena and Fernandez-Gaucherand (1999), Cavazos-Cadena and Hernández-Hernández (2005), Howard and Matheson (1972), Jaquette (1976), Kawai (1987), Mandl (1971), Sladký (2005),(2008),(2013), van Dijk and Sladký (2006), White (1988)). In the present paper we consider unichain Markov reward processes with finite state spaces and assume that the generated reward is evaluated by an exponential utility function. Using the Taylor expansion we present explicit formulae for calculating variance and higher central moments of the total reward generated by the Markov reward chain along with its asymptotic behavior and the growth rates if the considered time horizon tends to infinity.

Keywords Discrete-time Markov reward chains, exponential utility, moment generating functions, formulae for central moments.
JEL Classification C44, C61
AMS Classification 90C40

1 Formulation and Notation

In this note, we consider at discrete time points Markov reward process $X = \{X_n, n = 0, 1, \ldots\}$ with finite state space $\mathcal{I} = \{1, 2, \ldots, N\}$, matrix of transition probabilities $P = [p_{ij}]$ and transition reward matrix $R = [r_{ij}]$, i.e. reward r_{ij} is accrued to a transition from state i to state j. The symbol E_i denotes the expectation if $X_0 = i$; $\mathrm{P}(X_m = j)$ is the probability that X is in state j at time m. Moreover, I denotes an identity matrix and e is reserved for a unit column vector.

Recall that $P^* := \lim_{n\to\infty} \frac{1}{n}\sum_{k=0}^{n-1} P^k$ (with elements p^*_{ij}) exists, and if P is aperiodic then even $P^* = \lim_{k\to\infty} P^k$ and the convergence is geometrical. Moreover, if P is unichain, i.e. P contains a single class of recurrent states, then $p^*_{ij} = p^*_j$, i.e. limiting distribution is independent of the starting state (see e.g. Puterman (1994), Ross (1983)).

In what follows, the reward generated by the Markov chain X is evaluated by an exponential utility function, say $u^\gamma(\cdot)$, i.e. utility function with constant risk sensitivity $\gamma \in \mathbb{R}$, where

$$u^\gamma(\xi) := \begin{cases} \operatorname{sign}(\gamma)\, \exp(\gamma\xi) & \text{if } \gamma \neq 0 \\ \xi & \text{for } \gamma = 0. \end{cases} \tag{1}$$

Obviously, $u^\gamma(\cdot)$ is strictly increasing and convex, if $\gamma > 0$, for $\gamma = 0$ (the risk neutral case) $u^\gamma(\xi) = \xi$ is linear, if $\gamma < 0$ then $u^\gamma(\cdot)$ is concave (see e.g. Howard and Matheson (1972), Jaquette (1976), Sladký (2008),(2013)).

Observe that $U^{(\gamma)}(\xi) := \mathsf{E}\,\exp(\gamma\xi)$ is also the moment generating function of ξ provided the expectation is finite for $|\gamma| < h$, and some $h > 0$. As it is well-known (see e.g. Gut(2004), Th. 8.3) then

$$\mathsf{E}\,\exp[\gamma\xi]^k < \infty \text{ for all } k = 1, 2, \ldots \qquad \mathsf{E}\,\xi^n = \frac{\mathrm{d}^n}{\mathrm{d}\gamma^n} U^{(\gamma)}(\xi)|_{\gamma=0}, \quad (n = 1, 2, \ldots) \tag{2}$$

and the Taylor expansion around $\gamma = 0$ reads

$$U^{(\gamma)}(\xi) = 1 + \mathsf{E} \sum_{k=1}^{\infty} \frac{(\gamma\xi)^k}{k!} = 1 + \sum_{k=1}^{\infty} \frac{\gamma^k}{k!} \cdot \mathsf{E}\,\xi^k \quad \text{for} \quad |\gamma| < h. \tag{3}$$

Considering Markov reward chains, let

$\xi^{(n)} = \sum_{k=0}^{n-1} r_{X_k, X_{k+1}}$ be the (random) total reward received in the n next transitions of the considered Markov chain X.

Supposing that $X_0 = i$, on taking expectation we have for the first and second moment of $\xi^{(n)}$

$$V_i(n) := \mathsf{E}_i\,(\xi^{(n)}) = \mathsf{E}_i \sum_{k=0}^{n-1} r_{X_k, X_{k+1}}, \qquad S_i(n) := \mathsf{E}_i\,(\xi^{(n)})^2 = \mathsf{E}_i\,(\sum_{k=0}^{n-1} r_{X_k, X_{k+1}})^2 \tag{4}$$

hence the corresponding variance (i.e. the second central moment)

$$\sigma_i(n) := \mathsf{E}_i\,[\xi^{(n)} - V_i(n)]^2 = S_i(n) - [V_i(n)]^2. \tag{5}$$

Similarly, if the chain starts in state i for the expected utility we have

$$U_i(\gamma, n) := \mathsf{E}_i\,[\exp(\gamma\xi^{(n)})] = \mathsf{E}_i \exp[\gamma\,(r_{i,X_1} + \xi^{(1,n)}_{X_1})], \tag{6}$$

where $\xi^{(m,n)}_{X_m}$ (for $m < n$) is the reward obtained in the interval $[m, n]$ starting with state X_m.

In what follows let, $U(\gamma, n)$ be the (column) vector of expected utilities with elements $U_i(\gamma, n)$. Conditioning in (6) on X_1, from (6) we immediately get the recurrence formula

$$U_i(\gamma, n+1) = \sum_{j \in \mathcal{I}} p_{ij} \cdot \mathrm{e}^{\gamma r_{ij}} \cdot U_j(\gamma, n) = \sum_{j \in \mathcal{I}} q_{ij} \cdot U_j(\gamma, n) \quad \text{where} \quad U_j(\gamma, 0) = 1 \tag{7}$$

$$U(\gamma, n+1) = Q \cdot U(\gamma, n) \quad \text{with } U(\gamma, 0) = e, \text{ where } Q = [q_{ij}] = P \otimes R \tag{8}$$

the symbol $\otimes$ is used for the Hadamard product of matrices, i.e. $q_{ij} := p_{ij} \cdot \mathrm{e}^{\gamma r_{ij}}$

The paper is organized as follows. Section 2 contains basic formulae for calculating higher and higher central moments. Explicit formulae for higher central moments can be found in section 3, growth rates of central moments are discussed in section 4. Conclusions are made in the last section.

2 Exponential Utility and Higher Moments

Recall that by (7) $U_i(\gamma, n) = \mathsf{E}_i\,[\exp(\gamma\xi^{(n)})]$ is also the moment generating function of $\xi^{(n)}$. Hence (cf. (2)) for some $h > 0$ and any $|\gamma| < h$

$\frac{\mathrm{d}}{\mathrm{d}\gamma}\mathsf{E}_i\,[\exp(\gamma\xi^{(n)})] = \mathsf{E}_i\,\xi^{(n)}[\exp(\gamma\xi^{(n)})]$, hence for $k = 0,1,2\ldots,\ n = 0,1,2\ldots$

$$M_i^{(k)}(n) := \mathsf{E}_i\left(\xi^{(n)}\right)^k = \frac{\mathrm{d}^k}{\mathrm{d}\gamma^k}\mathsf{E}_i\,[\exp(\gamma\xi^{(n)})]|_{\gamma=0} \ \text{ is the } \ k\text{th moment of } \ \xi^{(n)} \tag{9}$$

and (cf. (3)) the Taylor expansion around $\gamma = 0$ reads

$$U_i(\gamma,n) = 1 + \sum_{k=1}^{\infty}\frac{\gamma^k}{k!}M_i^{(k)}(n) \ \text{ for } \ |\gamma| < h. \tag{10}$$

This along with the identities (6), (7) will be extremely useful for finding explicit formulas of moments of $\xi^{(n)}$. In particular, since for $|\gamma| < h$

$$\mathrm{e}^{\gamma r_{ij}} = 1 + \sum_{k=1}^{\infty}\frac{\gamma^k}{k!}[r_{ij}]^k$$

from (7),(10) we immediately get

$$1 + \sum_{k=1}^{\infty}\frac{\gamma^k}{k!}M_i^{(k)}(n+1) = \sum_{j\in\mathcal{I}}p_{ij}\left(1 + \sum_{k=1}^{\infty}\frac{\gamma^k}{k!}[r_{ij}]^k\right)\left(1 + \sum_{k=1}^{\infty}\frac{\gamma^k}{k!}M_j^{(k)}(n)\right). \tag{11}$$

Similarly on introducing the moment generating function for the central moments of $\xi^{(n)}$ by

$$\tilde{U}_i(\gamma,n) := \mathsf{E}_i\,[\exp(\gamma(\xi^{(n)} - \mathsf{E}_i\,\xi^{(n)})] \quad \text{for all} \ \ i \in \mathcal{I} \tag{12}$$

for the kth central moment of $\xi^{(n)}$ we have

$$\widetilde{M}_i^{(k)}(n) := \mathsf{E}_i\left[\xi^{(n)} - \mathsf{E}_i\,\xi^{(n)}\right]^k = \frac{\mathrm{d}^k}{\mathrm{d}\gamma^k}\mathsf{E}_i\,[\exp(\gamma(\xi^{(n)} - \mathsf{E}_i\,\xi^{(n)})]|_{\gamma=0} \tag{13}$$

and the Taylor expansion around $\gamma = 0$ for $|\gamma| < h$ reads

$$\tilde{U}_i(\gamma,n) = 1 + \sum_{k=1}^{\infty}\frac{\gamma^k}{k!}\cdot\widetilde{M}_i^{(k)}(n). \tag{14}$$

For what follows, let for real g, w_i's ($i \in \mathcal{I}$)

$$\tilde{\varphi}_{ij}(w,g) := r_{ij} - g + w_j - w_i \ \text{ where } \ c = \max_{i\in\mathcal{I}}|w_i|.$$

Then $r_{X_k,X_{k+1}} = \tilde{\varphi}_{X_k,X_{k+1}}(w,g) + g + w_{k+1} - w_k$ and from the first part of (7) we arrive at we arrive at

$$V_i(n) = ng + w_i + \mathsf{E}_i\,[\sum_{k=0}^{n-1}\varphi_{X_k,X_{k+1}}(w,g) - w^{\gamma}_{X_n}]. \tag{15}$$

It is well-known that for unichain transition matrix P w_i's and g can be selected such that if

$$\sum_{i\in\mathcal{I}}p_{ij}\varphi_{ij}(w,g) = 0 \ \text{ for all } \ i \in \mathcal{I} \ \text{ then } \ V_i(n) = ng + w_i + \mathsf{E}_i\,w_{X_n} = 0$$

Similarly, from (7) we conclude that

$$U_i(\gamma,n) = \mathsf{E}_i\,\mathrm{e}^{\gamma\sum\limits_{k=0}^{n-1}r_{X_k,X_{k+1}}} = \mathrm{e}^{\gamma[ng+w_i^{\gamma}]} \times \mathsf{E}_i\,\mathrm{e}^{\gamma[\sum\limits_{k=0}^{n-1}\tilde{\varphi}_{X_k,X_{k+1}}(w^{\gamma},g)-w^{\gamma}_{X_n}]}. \tag{16}$$

Observe that the first term on the RHS of (15) is non-random and hence

$$\mathsf{E}_i\, \mathrm{e}^{\gamma[\sum\limits_{k=0}^{n-1} \tilde{\varphi}_{X_k,X_{k+1}}(w^\gamma,g)-c]} \le \frac{U_i(\gamma,n)}{\mathrm{e}^{\gamma[ng+w_i^\gamma]}} \le \mathsf{E}_i\, \mathrm{e}^{\gamma[\sum\limits_{k=0}^{n-1} \tilde{\varphi}_{X_k,X_{k+1}}(w^\gamma,g)+c]} \tag{17}$$

If w_i^γ's and g are selected such that for any $i \in \mathcal{I}$

$$\sum_{j\in\mathcal{I}} p_{ij}\, \mathrm{e}^{\gamma[r_{ij}-w_i^\gamma+w_j^\gamma-g]} = 1 \Longleftrightarrow \sum_{j\in\mathcal{I}} p_{ij}\, \mathrm{e}^{\gamma[r_{ij}+w_j^\gamma]} = \mathrm{e}^{\gamma[g+w_i]} \tag{18}$$

then for $X_0 = \ell$, $k = 0, 1, \dots$ $\mathsf{E}_\ell\,\{\exp[\gamma\tilde{\varphi}_{X_k,X_{k+1}}(w^\gamma, g)]|X_k = m\} = 1$.
Since for $v_i := \mathrm{e}^{\gamma w_i}$, $\rho := \mathrm{e}^g$ the RHS of the second equality of (18) can be also written as $\sum_{j\in\mathcal{I}} q_{ij}\, v_j = \rho v_i$, the (column) vector v (with elements v_i's) is the Perron eigenvector of a nonnegative matrix Q with elements $q_{ij} = p_{ij}\, \mathrm{e}^{\gamma r_{ij}}$ and ρ is the spectral radius of Q. It is well known that if Q is irreducible then the Perron eigenvector v can be selected strictly positive. Observe that Q is irreducible if the matrix P is irreducible. Moreover, v can be selected strictly positive if the matrix P is only unichain and the risk sensitive coefficient γ is sufficiently close to null (cf. Gantmakher (1959)).

Since for suitably selected g, w_i's $\mathsf{E}\, \mathrm{e}^{\gamma\varphi_{X_{n-1},X_n}(w,g)+c}|X_{n-1} = i] = \mathrm{e}^{\gamma c}$ holds, then for the RHS of (18) we can conclude that

$$\mathsf{E}_i\, \mathrm{e}^{\gamma[\sum\limits_{k=0}^{n-1} \tilde{\varphi}_{X_k,X_{k+1}}(w,g)-c]} = \mathsf{E}_i\, \mathrm{e}^{\gamma[\sum\limits_{k=0}^{n-2} \tilde{\varphi}_{X_k,X_{k+1}}(w^\gamma,g)-c]}$$

and on iterating the above displayed formula we can conclude that

$$\mathsf{E}_i\, \mathrm{e}^{\gamma[\sum\limits_{k=0}^{n-1} \tilde{\varphi}_{X_k,X_{k+1}}(w^\gamma,g)-w']} = \mathsf{E}_i\, \mathrm{e}^{\gamma[-w']}.$$

Inserting into (18) we arrive at the following bounds on $U_i(\gamma, n)$

$$\mathrm{e}^{\gamma[ng+w_i^\gamma]-c} \le U_i(\gamma,n) \le \mathrm{e}^{\gamma[ng+w_i^\gamma]+c} \tag{19}$$

For the central moments similarly to (12), (13) we can conclude from (18), (19) that

$$\tilde{U}_i(\gamma,n) := \mathsf{E}_i\, \mathrm{e}^{\gamma[\xi^{(n)}-(ng-w_i+w_{X_n})]} = \sum_{j\in\mathcal{I}} p_{ij} \mathrm{e}^{\gamma(r_{ij}-g+w_i-w_j)} \tilde{U}_j(\gamma,n-1) \tag{20}$$

where

$$\tilde{U}_j(\gamma,n-1) = \mathsf{E}_j \mathrm{e}^{\gamma[\xi^{(1,n)}-(n-1)g+w_j-w_{X_n}]}.$$

In analogy to (11) we get

$$1+\sum_{k=1}^{\infty} \frac{\gamma^k}{k!} \widetilde{M}_i^{(k)}(n+1) = \sum_{j\in\mathcal{I}} p_{ij} \big(1+\sum_{k=1}^{\infty} \frac{\gamma^k}{k!} [r_{ij}-(g+w_i-w_j)]^k\big) \times \big(1+\sum_{k=1}^{\infty} \frac{\gamma^k}{k!} \widetilde{M}_j^{(k)}(n)\big). \tag{21}$$

3 Higher central moments: Explicit Formulas

Similarly as in the previous section our analysis based on (12),(13),(14) and (19) enables to generate recursively all central moments of $\xi^{(n)}$. Recall that

$$\widetilde{U}_i(\gamma,n+1) = \sum_{j\in\mathcal{I}} p_{ij}\, \mathrm{e}^{\gamma\,[r_{ij}-(g+w_i-w_j)]} \widetilde{U}_j(\gamma,n), \qquad \widetilde{U}_j(\gamma,n) = 1+\sum_{k=1}^{\infty} \frac{\gamma^k}{k!} \widetilde{M}_j^{(k)}(n)$$

$$\mathrm{e}^{\gamma\,[r_{ij}-(g+w_i-w_j)]} = 1+\sum_{k=1}^{\infty} \frac{\gamma^k}{k!} [r_{ij}-(g+w_i-w_j)]^k$$

By comparing in (20) the terms γ^k $(k = 1, 2, \ldots)$ we obtain the following recursive formulas for the central moments (obviously, the first central moment $\widetilde{M}_i^{(1)}(n) \equiv 0$ for all n).

In particular,

$$\text{For } k = 1 : \widetilde{M}_i^{(1)}(n+1) = \sum_{j\in\mathcal{I}} p_{ij}\, \widetilde{M}_j^{(1)}(n) \Rightarrow \widetilde{M}_j^{(1)}(n) = 0. \tag{22}$$

$$\text{For } k = 2 : \widetilde{M}_i^{(2)}(n+1) = \sum_{j\in\mathcal{I}} p_{ij}\, [(r_{ij} + w_j) - (g + w_i)]^2 + \sum_{j\in\mathcal{I}} p_{ij}\, \widetilde{M}_j^{(2)}(n). \tag{23}$$

$$\begin{aligned}\text{For } k = 3 : \widetilde{M}_i^{(3)}(n+1) &= \sum_{j\in\mathcal{I}} p_{ij}\, [(r_{ij} + w_j) - (g + w_i)]^3 \\ &+3 \sum_{j\in\mathcal{I}} p_{ij}\, [(r_{ij} + w_j) - (g + w_i)]\, \widetilde{M}_j^{(2)}(n) + \sum_{j\in\mathcal{I}} p_{ij}\, \widetilde{M}_j^{(3)}(n).\end{aligned} \tag{24}$$

$$\begin{aligned}\text{For } k = 4 : \widetilde{M}_i^{(4)}(n+1) &= \sum_{j\in\mathcal{I}} p_{ij}\, [(r_{ij} + w_j) - (g + w_i)]^4 \\ &+ 6 \sum_{j\in\mathcal{I}} p_{ij}\, [(r_{ij} + w_j) - (g + w_i)]^2\, \widetilde{M}_j^{(2)}(n) \\ &+ 4 \sum_{j\in\mathcal{I}} p_{ij}\, [(r_{ij} + w_j) - (g + w_i)]\, \widetilde{M}_j^{(3)}(n) \; + \; \sum_{j\in\mathcal{I}} p_{ij}, \widetilde{M}_j^{(4)}(n)\end{aligned} \tag{25}$$

In general:

$$\begin{aligned}\widetilde{M}_i^{(s)}(n+1) &= \sum_{j\in\mathcal{I}} p_{ij}\; \{[(r_{ij} + w_j) - (g + w_i)]^s\} \\ &+ \sum_{j\in\mathcal{I}} p_{ij} \left\{ \sum_{k=1}^{s-1} \binom{s}{k} [(r_{ij} + w_j) - (g + w_i)]^k\, \widetilde{M}_j^{(s-k)}(n) \right\} + \sum_{j\in\mathcal{I}} p_{ij}\, \widetilde{M}_j^{(s)}(n)\end{aligned} \tag{26}$$

that can be also written as

$$\widetilde{M}_i^{(s)}(n+1) = \sum_{k=0}^{s} \binom{s}{k} \sum_{j\in\mathcal{I}} p_{ij} \left\{ [(r_{ij} + w_j) - (g + w_i)]^k\, \widetilde{M}_j^{(s-k)}(n) \right\} \tag{27}$$

4 Growth rates of the central moments

To begin with, let us recall the result for the growth rate of the variance of the total reward. In virtue of (23) we immediately conclude that the variance (i.e. the central second moment) of the total reward grows linearly over time and the growth rate $g^{(2)}$ of $\widetilde{M}_i^{(2)}(n)$ in (24) can be found as a solution of

$$\widetilde{M}_i^{(2)}(n) = ng^{(2)} + w_i^{(2)} \qquad \text{where} \tag{28}$$

$$g^{(2)} + w_i^{(2)} = s_i^{(2)} + \sum_{j\in\mathcal{I}} p_{ij} w_j^{(2)}, \quad s_i^{(2)} = \sum_{j\in\mathcal{I}} p_{ij} [(r_{ij} + w_j) - (g + w_i)]^2. \tag{29}$$

To establish the growth rate of $\widetilde{M}_i^{(3)}(n)$, it suffices to insert into (24) from (23). Since $\sum_{j\in\mathcal{I}} p_{ij}\,[(r_{ij}+w_j)-(g+w_i)]\,g^{(2)}=0$ we can conclude that

$$\sum_{j\in\mathcal{I}} p_{ij}\,[(r_{ij}+w_j)-(g+w_i)]\,(ng^{(2)}+w_j^{(2)}) = \sum_{j\in\mathcal{I}} p_{ij}\,[(r_{ij}+w_j)-(g+w_i)]\,w_j^{(2)}.$$

On inserting into (23) we have

$$\widetilde{M}_i^{(3)}(n+1) = \sum_{j\in\mathcal{I}} p_{ij}\,[(r_{ij}+w_j)-(g+w_i)]^3 + 3\sum_{j\in\mathcal{I}} p_{ij}\,[(r_{ij}+w_j)-(g+w_i)]\,w_j^{(2)} + \sum_{j\in\mathcal{I}} p_{ij}\,\widetilde{M}_j^{(3)}(n). \qquad (30)$$

Hence using the same arguments as for the second central moment we can conclude that

$$\widetilde{M}_i^{(3)}(n) = ng^{(3)}+w_i^{(3)} \quad \text{where} \quad g^{(3)}+w_i^{(3)} = s_i^{(3)} + \sum_{j\in\mathcal{I}} p_{ij}w_j^{(3)} \qquad (31)$$

$$s_i^{(3)} = \sum_{j\in\mathcal{I}} p_{ij}\left\{[(r_{ij}+w_j)-(g+w_i)]^3 + 3\,[(r_{ij}+w_j)-(g+w_i)]\,w_j^{(2)}\right\}. \qquad (32)$$

Conclusion: The growth of the second and the third central moments of the total reward is linear over time.

Unfortunately, this approach cannot be directly applied for finding formulas for the higher central moments of exponential utility functions as the following analysis can show. In particular, to establish the growth rate of $\widetilde{M}_i^{(4)}(n)$, we insert into (25) from (24) and (23). After some algebra, since $\sum_{j\in\mathcal{I}} p_{ij}\,[(r_{ij}+w_j)-(g+w_i)]\,g^{(3)}=0$, we arrive at

$$\widetilde{M}_i^{(4)}(n+1) = \sum_{j\in\mathcal{I}} p_{ij}\,[(r_{ij}+w_j)-(g+w_i)]^4 + 6\sum_{j\in\mathcal{I}} p_{ij}\,[(r_{ij}+w_j)-(g+w_i)]^2\,(ng^{(2)}+w_j^{(2)}) + 4\sum_{j\in\mathcal{I}} p_{ij}\,[(r_{ij}+w_j)-(g+w_i)]\,w_j^{(3)} + \sum_{j\in\mathcal{I}} p_{ij}\,\widetilde{M}_j^{(4)}(n). \qquad (33)$$

Observe that the above equations (33) for the fourth central moments $\widetilde{M}_i^{(4)}(n)$ have the same structure as recursive equations (23),(24) for second a third central moments $\widetilde{M}_i^{(2)}(n)$, $\widetilde{M}_i^{(3)}(n)$. Unfortunately, the second term on the RHS of (33) contains also n, hence analogy with formulas for calculating average reward of a (unichain) Markov decision process cannot be used.

5 Conclusions

In the paper explicit formulae for calculating variance and higher central moments of the total reward generated by the Markov reward chain were obtained along with their growth rates if the considered time horizon tends to the infinity. In particular, it is shown that not only the variance (i.e. the second central moment) but also the skewness (i.e. the third central moment) grow linearly over time. Unfortunately, the linear growth rate does not hold for higher central moments.

Acknowledgements: This work was supported by the Czech Science Foundation under Grant 18-02739S.

References

[1] Cavazos-Cadena, R. and Fernandez-Gaucherand, F. (1999). Controlled Markov Chains with Risk-sensitive Criteria: Average Cost, Optimality Equations and Optimal Solutions. *Mathematical Methods of Operations Research* 43, pp. 121–139.

[2] Cavazos-Cadena, R. and Hernández-Hernández, D. (2005). A Characterization of the Optimal Risk-sensitive Average Cost Infinite Controlled Markov Chains. *Annals of Applied Probability* 15, pp. 175–212.

[3] Gantmakher, F. R. (1959). *The theory of matrices.* London: Chelsea.

[4] Gut, A. (2004). *Probability: a graduate course.* New York: Springer.

[5] Howard, R. A. and Matheson, J. (1972). Risk-sensitive Markov Decision Processes. *Management Science* 18, pp. 356–369.

[6] Jaquette, S. C. (1976). A Utility Criterion for Markov Decision Processes. *Management Science* 23, pp. 43–49.

[7] Kawai, H.(1987). A Variance Minimization Problem for a Markov Decision Process. *European Journal on Operational Research* 31, pp. 140–145.

[8] Mandl, P. (1971). On the Variance in Controlled Markov Chains. *Kybernetika* 7, pp. 1–12.

[9] Puterman, M. L. (1994). *Markov decision processes – discrete stochastic dynamic programming.* New York: Wiley.

[10] Ross, S. M. (1983). *Introduction to stochastic dynamic programming.* New York: Academic Press.

[11] Sladký, K. (2005). On Mean Reward Variance in Semi-Markov Processes. *Mathematical Methods of Operations Research* 62, pp. 387–397.

[12] Sladký, K. (2008). Growth Rates and Average Optimality in Risk-sensitive Markov Decision Chains. *Kybernetika* 44, pp. 206–217.

[13] Sladký, K.(2013). Risk-sensitive and Mean Variance Optimality in Markov Decision Processes. *Acta Oeconomica Pragensia* 7, pp. 146–161.

[14] van Dijk, N.M. and Sladký, K. (2006). On the Total Reward Variance for Continuous-time Markov Reward Chains. *Journal of Applied Probability* 43, pp. 1044-1052.

[15] White, D. J. (1988). Mean, Variance and Probability Criteria in Finite Markov Decision Processes: A Review. *Journal of Optimization Theory and Applications* 6, pp. 1–29.

Author's address
Ing. Karel Sladký, CSc.
Institute of Information Theory and Automation
Academy of Sciences of the Czech Republic
Department of Econometrics
Pod Vodárenskou věží 4, 182 08 Praha 8
Czech Republic
e-mail: sladky@utia.cas.cz