Separable Utility Functions in Dynamic Economic Models

Karel Sladký[1]

Abstract. In this note we study properties of utility functions suitable for performance evaluation of dynamic economic models under uncertainty. At first, we summarize basic properties of utility functions, at second we show how exponential utility functions can be employed in dynamic models where not only expectation but also the risk are considered. Special attention is focused on properties of the expected utility and the corresponding certainty equivalents if the stream of obtained rewards is governed by Markov dependence and evaluated by exponential utility functions.

Keywords: Utility functions, decision under uncertainty, dynamic economic models, Markov reward chains, exponential utility functions, certainty equivalent

JEL classification: C44,D81
AMS classification: 91B16,90C40

1 Economic decisions and utility functions

Economic decisions are usually based on the outcome, say ξ, as viewed by the decision maker. Usually the decision maker has no complete information on the problem, his decisions are made under some kind of uncertainty. In general, the decision under uncertainty in its most simple form consists of three nonempty sets $\mathcal{I}$, $\mathcal{A}$ and $\mathcal{X}$ and a function $f : \mathcal{I} \times \mathcal{A} \mapsto \mathcal{X}$. In particular, $\mathcal{I} = \{1, 2, \dots, N\}$ characterizes the uncertainty of the problem, elements of the set $\mathcal{I}$ are called states of the considered system (sometimes called states of nature or of the world); $\mathcal{A} = \{1, 2, \dots, K\}$ is the set of possible decisions (or actions) and $\mathcal{X}$ is the set of outcomes of the decision problem equipped with a complete and transitive relation $\preceq$ on $\mathcal{X}$. As concerns the relation $\preceq$ that determines decision maker's preference among elements of the set of outcomes (and also preference among actions). In particular, if for $\xi_1, \xi_2 \in \mathcal{X}$ (resp. $a_1, a_2 \in \mathcal{A}$) it holds $\xi_1 \preceq \xi_2$ (resp. $a_1 \preceq a_2$) it means that outcome ξ_2 (resp. decision a_2) is at least as preferable as outcome ξ_1 (resp. decision a_1). By completeness of the relation we mean that every two elements of $\mathcal{X}$ (resp. of $\mathcal{A}$) are related, i.e. given any $\xi_1, \xi_2 \in \mathcal{X}$ there are three possibilities: either $\xi_1 \preceq \xi_2$ but not $\xi_2 \preceq \xi_1$, then we write $\xi_1 \prec \xi_2$; or $\xi_2 \preceq \xi_1$ but not $\xi_1 \preceq \xi_2$, then we write $\xi_2 \prec \xi_1$; or both $\xi_1 \preceq \xi_2$ and $\xi_2 \preceq \xi_1$, then we write $\xi_2 \sim \xi_1$. By transitivity we mean that $\xi_1 \preceq \xi_2$ and $\xi_2 \preceq \xi_3$ implies $\xi_1 \preceq \xi_3$ for any three elements $\xi_1, \xi_2, \xi_3 \in \mathcal{X}$; the same also holds for any three elements of the action set $\mathcal{A}$.

Furthermore, in many decision problems on choosing action (decision) $a \in \mathcal{A}$ the outcome $\xi_j \in \mathcal{X}$ occurs only with (individual's subjective) probability p_j^a for $j = 1, 2, \dots, N$ (where $\sum_{j=1}^{N} p_j^a = 1$), which is familiar to the decision maker (stochastic model). In this case we speak about lottery or prospect, and let $\mathcal{Y}$ with generic y be the set of all lotteries or all probability distribution on $\mathcal{Y}$. Obviously, if the decision maker had a complete ranking of all lotteries on the set of outcomes, then he could obtain a complete ranking of all decisions in $\mathcal{A}$. To this end the decision maker can replace condition on complete ordering of the set $\mathcal{X}$ by complete ordering of the set $\mathcal{Y}$ (see Axiom 1). Moreover, under some other technical assumptions (Axioms 2,3 specified in the sequel), ordering of decision may be expressed by numerical function, called the utility function.

Axiom 1. (Preference Ordering) The decision maker has a preference ordering defined on $\mathcal{Y}$ which is a transitive and complete ordering.

[1]Institute of Information Theory and Automation of the AS CR, Department of Econometrics, Pod Vodárenskou věží 4, 182 08 Praha 8, Czech Republic, e-mail: sladky@utia.cas.cz

Axiom 2. **(Continuity)** If for $y_1, y_2, y_3 \in \mathcal{Y}$, $y_1 \preceq y_2 \preceq y_3$, there exists an $\alpha \in [0,1]$ such that $\alpha y_1 + (1-\alpha)y_3 \sim y_2$.

Axiom 3. **(Independence)** If for $y_1, y_2 \in \mathcal{Y}$, $y_1 \preceq y_2$, then for all $\alpha \in (0,1]$ and all $y \in \mathcal{Y}$

$$\begin{array}{llll} \text{if} & y_1 \prec y_2 & \text{then also} & \alpha y_1 + (1-\alpha)y \preceq y_2 + (1-\alpha)y \\ \text{if} & y_1 \sim y_2 & \text{then also} & \alpha y_1 + (1-\alpha)y \sim y_2 + (1-\alpha)y \end{array}$$

Theorem. Under Axioms 1–3 there exists a real-valued function $u : \mathcal{X} \mapsto \mathbb{R}$, called utility function, such that for all $y_1, y_2 \in \mathcal{Y}$

$$y_1 \preceq y_2 \quad \text{if and only if} \quad \mathsf{E}^{y_1}[u(\xi)] \le \mathsf{E}^{y_2}[u(\xi)]$$

($\mathsf{E}^y[\cdot]$ is the expectation with respect to probability distribution induced by $y \in \mathcal{Y}$).

Furthermore, u is unique up to a positive linear transformation, i.e. if $\tilde{u}$ is another function with the above property, there exists a positive scalar α_1 and a scalar α_2, such that $\tilde{u}(\xi) = \alpha_1 u(\xi) + \alpha_2$.

For the *proof* of this fundamental theorem see e.g. [1], [2], [10].

In words: Economic decisions based on the outcome, say ξ, may be represented by an appropriate utility function, say $u(\xi)$ assigning a real number to each possible outcome ξ. Utility function $u(\xi)$ must be monotonically increasing, i.e. we assume that larger values of outcome are preferred, and unique up to a positive linear transformation. Furthermore, in economic models we assume that utility functions are concave.

In case of stochastic models outcome ξ is a random variable and we consider expectation of utilities assigned to (random) outcomes, i.e. the value $U(\xi) := \mathsf{E}u(\xi)$. Certain (or certainty) equivalent, say $Z(\xi)$, is then defined by $u(Z(\xi)) := \mathsf{E}u(\xi)$ (in words, certainty equivalent is the value, whose utility is the same as the expected utility of possible outcomes). Additional important concepts in the utility theory are that of the

coefficient of absolute risk aversion defined by $R^a(\xi) := -\frac{u''(\xi)}{u'(\xi)}$, along with

coefficient of relative risk aversion defined by $R^r(\xi) := \xi R^a(\xi) = -\xi \cdot \frac{u''(\xi)}{u'(\xi)}$.[1]

For handling real life models decision makers must be able to express $u(x)$ in a concrete form. Typical utility functions are:

- Linear function: $$u(x) = a + bx \quad \text{where} \quad b > 0$$
- Quadratic function: $$u(x) = a + bx - cx^2 \quad \text{where} \quad b > 0,\ c > 0.$$
- Logarithmic function: $$u(x) = a + b\ln(x+c) \quad \text{where} \quad b > 0,\ c \ge 0.$$
- Fractional function: $$u(x) = a - \frac{1}{x+b} \quad \text{where} \quad b > 0,\ c > 0$$
- The function: $$u(x) = \begin{cases} x^{1-a} & \text{for} \quad 0 < a < 1 \\ \ln x & \text{for} \quad a = 1 \\ -x^{1-a} & \text{for} \quad a > 1 \end{cases}$$

  Observe that then $R^r(x) = a$, i.e. coefficient of relative risk aversion is constant; this utility function belong to the CRRA (Constant Relative Risk Aversion) utility functions.
- Exponential function: $$u(x) = -\mathrm{e}^{-ax} \text{ with } a > 0$$

Introducing the so-called risk aversion coefficient $\gamma \in \mathbb{R}$ exponential utility functions, as well as linear utility functions, assigned to a random outcome ξ can be also written in the following more compact form

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

[1] (here $u'(\xi) = \frac{\mathrm{d}u(\xi)}{\mathrm{d}\xi}$, $u''(\xi) = \frac{\mathrm{d}^2 u(\xi)}{\mathrm{d}\xi^2}$).

Observe that exponential utility function considered in (1) is separable what is very important for sequential decision problems, i.e. $u^\gamma(\xi^{(1)} + \xi^{(2)}) = \mathrm{sign}(\gamma)\, u^\gamma(\xi^{(1)}) \cdot u^\gamma(\xi^{(2)})$ for $\gamma \neq 0$ and if $\gamma = 0$ for the resulting linear utility function $u(x) = bx$ we have $u^\gamma(\xi^{(1)} + \xi^{(2)}) = u^\gamma(\xi^{(1)}) + u^\gamma(\xi^{(2)})$. Unfortunately, considering stochastic models, in contrast to exponential utility functions, linear utility functions cannot reflect variability-risk features of the problem. Obviously $u^\gamma(\cdot)$ is continuous and strictly increasing, and *convex* for $\gamma > 0$, so-called risk seeking case, and *concave* for $\gamma < 0$, so-called risk aversion case.

Furthermore, exponential utility functions

- are the most widely used non-linear utility functions, cf. [3],
- in most cases an appropriately chosen exponential utility function is a very good approximation for general utility function, cf. [6].

If exponential utility (1) is considered, then for the corresponding certainty equivalent $Z^\gamma(\xi)$ given by

$$u^\gamma(Z^\gamma(\xi)) = \mathsf{E}[(\mathrm{sign}\ \gamma)\ \exp(\gamma\xi)]$$

we have

$$Z^\gamma(\xi) = \begin{cases} \frac{1}{\gamma} \ln\{\mathsf{E}\,[\exp(\gamma\xi)]\}, & \text{if } \gamma \neq 0 \\ \mathsf{E}[\xi] & \text{for } \gamma = 0. \end{cases} \tag{2}$$

Observe that if ξ is constant then $Z^\gamma(\xi) = \xi$, if ξ is nonconstant then by Jensen's inequality

$$\begin{aligned} Z^\gamma(\xi) &> \mathsf{E}\xi \quad (\text{if } \gamma > 0, \text{ the risk seeking case}) \\ Z^\gamma(\xi) &< \mathsf{E}\xi \quad (\text{if } \gamma < 0, \text{ the risk averse case}) \\ Z^\gamma(\xi) &= \mathsf{E}\xi \quad (\text{if } \gamma = 0, \text{ the risk neutral case}) \end{aligned}$$

The following facts will be useful in the sequel:

1. For $U^{(\gamma)}(\xi) := \mathsf{E}u^\gamma(\xi)$, i.e. $U^{(\gamma)}(\xi) := \mathsf{E}\exp(\gamma\xi)$, the Taylor expansion around $\gamma = 0$ reads

$$U^{(\gamma)}(\xi) = 1 + \mathsf{E}\sum_{k=1}^{\infty} \frac{(\gamma\xi)^k}{k!} = 1 + \sum_{k=1}^{\infty} \frac{\gamma^k}{k!} \cdot \mathsf{E}\xi^k. \tag{3}$$

Observe that in (3) the first (resp. second) term of the Taylor expansion is equal to $\gamma\mathsf{E}\xi$ (resp. $\frac{1}{2}(\gamma^2)\mathsf{E}\xi^2$). In particular, if for random variables ξ, ζ with $\mathsf{E}\xi = \mathsf{E}\zeta$ it holds $\mathsf{E}\xi^2 < \mathsf{E}\zeta^2$ (or equivalently $\mathrm{var}\,\xi < \mathrm{var}\,\zeta$) and $\mathsf{E}\xi^k$ are uniformly bounded in k then there exists $\gamma_0 > 0$ such that $U^{(\gamma)}(\xi) < U^{(\gamma)}(\zeta)$ for any $\gamma \in (-\gamma_0, \gamma_0)$.

2. In economic models (see e.g. [1], [10]) we usually assume that the utility function $u(\cdot)$ is increasing (i.e. $u'(\cdot) > 0$), concave (i.e. $u''(\cdot) < 0$) with $u(0) = 0$ and $u'(0) < \infty$ (so called the Inada condition).

Since a positive linear transformation of the utility function $u^\gamma(\xi)$ preserves the original preferences (see the Theorem, cf. also [1],[10]) we shall also consider the utility functions

$$\bar{u}^\gamma(x) \;=\; 1 - \exp(\gamma x), \quad \text{where } \ \gamma < 0 \quad (\text{the risk averse case}) \tag{4}$$

$$\tilde{u}^\gamma(x) \;=\; \exp(\gamma x) - 1, \quad \text{where } \ \gamma > 0 \quad (\text{the risk seeking case}) \tag{5}$$

and the function $\bar{u}^\gamma(x)$ satisfies all above conditions imposed on a utility function in economy theory. Observe that the Taylor expansions of $\bar{u}^\gamma(x)$ and of $\tilde{u}^\gamma(x)$ read

$$\bar{u}^\gamma(x) = \sum_{k=1}^{\infty} (-1)^{k+1} \frac{|\gamma|^k}{k!} \cdot x^k, \ \text{ where } \gamma < 0, \quad \tilde{u}^\gamma(x) = \sum_{k=1}^{\infty} \frac{\gamma^k}{k!} \cdot x^k, \ \text{ where } \gamma > 0 \tag{6}$$

and if $x = \xi$ is a random variable for the expected utilities we have

$$\bar{U}^\gamma(\xi) \;:=\; \mathsf{E}\bar{u}^\gamma(\xi) = 1 - U^{(\gamma)}(\xi) = \sum_{k=1}^{\infty} (-1)^{k+1} \frac{|\gamma|^k}{k!} \cdot \mathsf{E}\xi^k \tag{7}$$

$$\tilde{U}^\gamma(\xi) \;:=\; \mathsf{E}\tilde{u}^\gamma(\xi) = U^{(\gamma)}(\xi) - 1 = \sum_{k=1}^{\infty} \frac{\gamma^k}{k!} \cdot \mathsf{E}\xi^k. \tag{8}$$

Illustrative example. Consider an individual that may repeatedly bet \$1 on the toss of a fair coin or not bet at all. f he bets and guesses correctly he wins \$2, if he does not guess correctly, he losses \$1 and if he decides not to bet he receives compensation \$1. Here the state space $\mathcal{I}$ consists of two elements H (head) and T (tail). The action set $\mathcal{A}$ consists of three elements $\mathcal{A} = \{a_0, a_1, a_2\}$ where decision a_0 is not to bet, decision a_1 is bet on heads and decision a_2 is bet on tails. Then the set of outcomes $\mathcal{X} = \{0, 1, 2\}$ and the values of the function $f : \mathcal{I} \times \mathcal{A} \mapsto \mathcal{X}$ are given by $f(H, a_0) = 1$, $f(T, a_0) = 1$, $f(H, a_1) = 2$, $f(T, a_1) = 0$, $f(T, a_2) = 2$, $f(H, a_2) = 0$. Moreover, we may consider also some kind of lottery in the decision process, in particular we may assume that the selected decision occurs in accordance with a given probability mechanism.
Now we need a ranking among decisions that is consistent in a well-defined sense with our ranking of outcomes. Moreover, according to Theorem the ranking should be determined by a numerical function $u(\cdot)$ that maps the set of decisions $\mathcal{A}$ to the set of real numbers such that $a_i \preceq a_j$ if and only if $u(a_i) \preceq u(a_j)$ for all $a_i, a_j \in \mathcal{A}$ and $i \neq j$. If $u(\cdot)$ is linear, i.e. $u(x) = x$ then $\mathsf{E}^{a_0} u(\xi) = 1$, $\mathsf{E}^{a_1} u(\xi) = \mathsf{E}^{a_2} u(\xi) = \frac{1}{2}(2 + 0) = 1$, hence then the resulting expectation of the linear utility function is independent of the selected decision, i.e. $a_0 \sim a_1 \sim a_2$. However, using the exponential utility function $u^\gamma(x)$ optimal decision depends on the value of the risk aversion coefficient $\gamma \neq 0$. In particular, again $\mathsf{E}^{a_0} u(\xi) = 1$, but $\mathsf{E}^{a_1} u(\xi) = \mathsf{E}^{a_2} u(\xi) = \frac{1}{2}(\mathrm{e}^{\gamma 2} + 1)$. Hence if $\gamma > 0$ then $a_0 \prec a_1 \sim a_2$ and if $\gamma < 0$ then $a_1 \sim a_2 \prec a_0$.

Up to now we have studied properties of utility functions for models of "static" (stochastic) systems where the uncertainty is represented by the decision maker's ignorance of the "current state" of the system along with possibly probabilistic behavior of the outcome. In the sequel we focus attention on systems that develop over time and the decision maker need not have complete information of the state of the system and where additional decision can be taken if the individual guesses correctly.

2 Separable utility functions in stochastic dynamic models

2.1 Complete information on state

Consider a family of models for decision under uncertainty formulated in Section 1 specified by nonempty sets $\mathcal{I}$, $\mathcal{A}$ along with a family of $\mathcal{X}^{(i)}$ (for $= 1, 2, \dots N$) with individual probabilities p_{ij}^a for $i, j = 1, 2, \dots, N; a = 1, 2, \dots, K$ (of course, $\sum_{j=1}^N p_{ij}^k = 1$), familiar to the decision maker along with his (or her) knowledge of the current state of the system. This represents a Markov decision chain $X = \{X_n, n = 0, 1, \dots\}$ with finite state space $\mathcal{I} = \{1, \dots, N\}$, finite set $\mathcal{A}_i = \{1, 2, \dots, K\}$ of possible decisions (actions) in state $i \in \mathcal{I}$ and the following transition and reward structure:

p_{ij}^a : transition probability from $i \to j$; r_{ij}^a : one-stage reward for a transition from $i \to j$

Policy controlling the chain is a rule how to select actions in each state. In this note, we restrict on stationary policies, i.e. the rules selecting actions only with respect to the current state of the Markov chain X. Then a policy, say π, is determined by some decision vector f whose ith element $f_i \in \mathcal{A}$ identifies the action taken if the chain X is in state $X_n = i$; hence also the transition probability matrix $\boldsymbol{P}(f)$ of the Markov decision chain. Observe that the ith row of $\boldsymbol{P}(f)$ has elements $p_{i1}^{f_i}, \dots, p_{iN}^{f_i}$ and that $\boldsymbol{P}^*(f) = \lim_{n\to\infty} n^{-1} \sum_{k=0}^{n-1} [\boldsymbol{P}(f)]^k$ exists. In what follows, $\boldsymbol{R}(f) = [r_{ij}^{f_i}]$ is the transition reward matrix, i.e. $\boldsymbol{R}(f)$ is an $N \times N$ matrix of one-stage rewards (for details see e.g. [7]).

If the chain starts in state i and policy $\pi \sim (f)$ is followed, let $\xi_i^{(n)}(\pi)$ (abbreviated as $\xi^{(n)}$) be the total reward received in the n next transition of the Markov chain X and $\xi_{X_m}^{(m,n)}(\pi)$ be the total reward received in the $n-m$ next transition if the Markov chain was after m first transition in state X_m. Then for the expected (exponential) utility $U_i^\pi(\gamma, n)$, the certainty equivalent $Z_i^\pi(\gamma, n)$ and its mean value $J_i^\pi(\gamma)$ we have

$$U_i^\pi(\gamma, n) := \mathsf{E}_i^\pi[\exp(\gamma \xi^{(n)})] = \mathsf{E}_i^\pi \exp[\gamma\, (r_{i,X_1} + \xi_{X_1}^{(1,n)})], \tag{9}$$

$$Z_i^\pi(\gamma, n) := \frac{1}{\gamma} \ln\{\mathsf{E}_i^\pi\, [\exp(\gamma \xi^{(n)})]\} \qquad \text{for } \gamma \neq 0, \tag{10}$$

$$J_i^\pi(\gamma) := \lim_{n\to\infty} \frac{1}{n} Z_i^\pi(\gamma, n) \tag{11}$$

and hence for the expectation of the utility functions $\bar{u}^\gamma(\xi^{(n)})$ and $\tilde{u}^\gamma(\xi^{(n)})$ we have (cf. (8),(8))

$$\bar{U}_i^\pi(\gamma, n) := 1 - U_i^\pi(\gamma, n), \qquad \tilde{U}_i^\pi(\gamma, n) := U_i^\pi(\gamma, n) - 1. \tag{12}$$

Conditioning in (9) on X_1, since policy $\pi \sim (f)$ is stationary, from (9) we immediately get the recurrence formula

$$U_i^\pi(\gamma, n+1) \;=\; \sum_{j\in\mathcal{I}} p_{ij}^{f_i} \cdot \mathrm{e}^{\gamma r_{ij}} \cdot U_j^\pi(\gamma, n) = \sum_{j\in\mathcal{I}} q_{ij}^{f_i} \cdot U_j^\pi(\gamma, n) \quad \text{with} \;\; U_i^\pi(\gamma, 0) = 1 \tag{13}$$

or in vector notation and by iterating

$$\boldsymbol{U}^\pi(\gamma, n+1) \;=\; \boldsymbol{Q}(f) \cdot \boldsymbol{U}^\pi(\gamma, n) = (\boldsymbol{Q}(f))^n \cdot \boldsymbol{e} \qquad \text{with} \;\; \boldsymbol{U}^\pi(\gamma, 0) = \boldsymbol{e}, \tag{14}$$

where $\boldsymbol{Q}(f) = [q_{ij}^{f_i}]$ with $q_{ij}^{f_i} := p_{ij}^{f_i} \cdot \mathrm{e}^{\gamma r_{ij}}$, $\boldsymbol{U}^\pi(\gamma, n)$ is the vector of expected utilities with elements $U_i^\pi(\gamma, n)$ and $\boldsymbol{e}$ is a unit (column) vector.

Observe that $\boldsymbol{Q}(f)$ is a nonnegative matrix, and by the Perron–Frobenius theorem (cf. [4]) the spectral radius $\rho(f)$ of $\boldsymbol{Q}(f)$ is equal to the maximum positive eigenvalue of $\boldsymbol{Q}(f)$. Moreover, if $\boldsymbol{Q}(f)$ is irreducible (i.e. if and only if $\boldsymbol{P}(f)$ is irreducible) the corresponding (right) eigenvector $\boldsymbol{v}(f)$ can be selected strictly positive, i.e.

$$\rho(f)\, \boldsymbol{v}(f) = \boldsymbol{Q}(f) \cdot \boldsymbol{v}(f) \qquad \text{with} \quad \boldsymbol{v}(f) > 0. \tag{15}$$

Moreover, under the above irreducibility condition it can be shown (cf. e.g. [5], [9]) that there exists decision vector $f^* \in \mathcal{A}$ such that

$$\boldsymbol{Q}(f) \cdot \boldsymbol{v}(f^*) \;\le\; \rho(f^*)\, \boldsymbol{v}(f^*) = \boldsymbol{Q}(f^*) \cdot \boldsymbol{v}(f^*), \tag{16}$$

$$\rho(f) \;\le\; \rho(f^*) \equiv \rho^* \qquad \text{for all} \;\; f \in \mathcal{A} \tag{17}$$

and decision vector $\hat{f} \in \mathcal{A}$ such that

$$\boldsymbol{Q}(f) \cdot \boldsymbol{v}(\hat{f}) \;\ge\; \rho(\hat{f})\, \boldsymbol{v}(\hat{f}) = \boldsymbol{Q}(\hat{f}) \cdot \boldsymbol{v}(\hat{f}) \tag{18}$$

$$\rho(f) \;\ge\; \rho(\hat{f}) \equiv \hat{\rho} \qquad \text{for all} \;\; f \in \mathcal{A}. \tag{19}$$

In words, $\rho(f^*) \equiv \rho^*$ (resp. $\rho(\hat{f}) \equiv \hat{\rho}$) is the maximum (resp. minimum) possible positive eigenvalue of $\boldsymbol{Q}(f)$ over all $f \in \mathcal{A}$.

If the Perron eigenvectors $\boldsymbol{v}(f^*) = \boldsymbol{v}^*$, $\boldsymbol{v}(\hat{f}) = \hat{\boldsymbol{v}}$ are strictly positive, there exist positive numbers $\alpha_1 < \alpha_2$ such that $\alpha_1 \hat{\boldsymbol{v}} \le \boldsymbol{e} \le \alpha_2 \boldsymbol{v}^*$ and hence by (16), (18) and by (10),(11)

$$\alpha_1 \hat{\rho}^n \hat{\boldsymbol{v}} \le \boldsymbol{U}^\pi(\gamma, n) \le \alpha_2 (\rho^*)^n \boldsymbol{v}^* \tag{20}$$

$$n \ln(\hat{\rho}) + \ln(\alpha_1 \hat{v}_i) \le \gamma Z_i^\pi(\gamma, n) \le n \ln(\rho^*) + \ln(\alpha_2 v_i^*) \tag{21}$$

$$\gamma^{-1} \ln(\hat{\rho}) \le J_i^\pi(\gamma) \le \gamma^{-1} \ln(\rho^*) \tag{22}$$

From (20),(21),(22) we can see that the asymptotic behavior of $\boldsymbol{U}^\pi(\gamma, n)$ heavily depends on ρ^*, $\hat{\rho}$, and that the maximum, resp. minimum, growth rate of each $U_i^\pi(\gamma, n)$ is independent of the starting state. Similarly, $J_i^\pi(\gamma)$ (mean value of the corresponding certainty equivalent $Z_i^\pi(n, \gamma)$ growing linearly in time) is independent of the starting state and bounded by $\ln(\hat{\rho})$ and by $\ln(\rho^*)$.

2.2 Incomplete information on state

In what follows we assume that the decision maker has no information of the current state of the system, but he knows current values of the obtained rewards. Moreover, he can also employ results concerning optimal policy obtained in subsection 2.1, i.e. the decision maker knows optimal actions in each state of the system. His knowledge of the system structure and optimal control policy along with information of current values of obtained rewards ("signalling information") may help him to selected optimal or suboptimal policy. In what follows we sketch how to handle such problems on examples slightly extending our illustrative example and reformulate it in terms of Markov decision chains (MDC).

Illustrative example: Formulation as MDC. Consider a Markov decision chain with state space $\mathcal{I} = \{H, T\}$, actions a_1, a_2 in each state and the following transition and reward structure:

$p_{HH}^{a_1} = \frac{1}{2}, r_{HH}^{a_1} = 2;\;\; p_{HT}^{a_1} = \frac{1}{2}, r_{TT}^{a_1} = 0;\quad p_{HH}^{a_2} = \frac{1}{2}, r_{HH}^{a_2} = 0;\;\; p_{HT}^{a_2} = \frac{1}{2}, r_{HT}^{a_2} = 2;$
$p_{TT}^{a_1} = \frac{1}{2},\;\; r_{TT}^{a_1} = 0;\;\; p_{TH}^{a_1} = \frac{1}{2},\;\; r_{TH}^{a_1} = 2;\quad p_{TH}^{a_2} = \frac{1}{2}, r_{TH}^{a_2} = 0;\;\; p_{TT}^{a_2} = \frac{1}{2}, r_{TT}^{a_2} = 2.$

Recall that a_1, a_2 is bet on heads, tails respectively; we ignore decision a_0 not to bet. In what follows $q^a_{ij} = p^a_{ij}\exp(\gamma r^a_{ij})$ for $a = a_1, a_2$ and $i, j = H, T$, in particular, $q^a_{ij} = \frac{1}{2}\exp(\gamma 2)$ or $q^a_{ij} = \frac{1}{2}\cdot 1$.
Hence each row of the resulting 2×2 matrix $\boldsymbol{Q}(\cdot) = [q^k_{ij}]$ contains a single element $\frac{1}{2}\exp(\gamma 2)$ and $\frac{1}{2}$. Hence the spectral radius of each $\boldsymbol{Q}(\cdot)$ equals $\frac{1}{2}[\exp(\gamma 2) + 1]$ and constant vector is the corresponding right eigenvector. Since decision a_0 (not to bet) brings unit reward, if the risk aversion coefficient γ is positive, the decision maker should prefer betting on heads or on tails, if γ is negative optimal decision is not to bet.

Extended illustrative example: Formulation as MDC. Suppose that in the considered Illustrative Example an individual can extend options after his betting on heads and guessing correctly. The additional action is to bet \$1 on heads and toss of an unfair coin (probability of head $\frac{1}{3}$, probability of tail $\frac{2}{3}$). If he guesses correctly receives \$3 and can repeat such bet.
To this end consider a Markov decision chain with state space $\mathcal{I} = \{H, T, \bar{H}\}$ and slightly modified transition and reward structure of the previous example by replacing transition from state H if action a_1 (instead of $p^{a_1}_{HH} = \frac{1}{2}, r^{a_1}_{HH} = 2$) by $p^{a_1}_{H\bar{H}} = \frac{1}{2}, r^{a_1}_{H\bar{H}} = 2$ and define transitions and rewards in state $\bar{H}$ as:

$p^{a_1}_{\bar{H}\bar{H}} = \frac{1}{2}, r^{a_1}_{\bar{H}\bar{H}} = 2;\ \ p^{a_1}_{\bar{H}T} = \frac{1}{2}, r^{a_1}_{\bar{H}T} = 0;\ \ p^{a_2}_{\bar{H}H} = \frac{1}{2}, r^{a_2}_{\bar{H}H} = 0;\ p^{a_2}_{\bar{H}T} = \frac{1}{2}, r^{a_2}_{\bar{H}T} = 2;$

$p^{a_3}_{\bar{H}\bar{H}} = \frac{1}{3},\ r^{a_3}_{\bar{H}\bar{H}} = 3;\ p^{a_3}_{\bar{H}T} = \frac{2}{3}, r^{a_3}_{\bar{H}T} = 0.$

Obviously, if action a_3 is not selected in state $\bar{H}$ the problem can be treated as the previous one. To this end we focus attention on policies selecting decision a_3 and starting with decision a_1. Then the spectral radius of the corresponding 3×3 matrix $\boldsymbol{Q}(\cdot) = \begin{bmatrix} 0 & \frac{1}{2} & \frac{1}{2}\mathrm{e}^{\gamma 2} \\ \frac{1}{2} & \frac{1}{2}\mathrm{e}^{\gamma 2} & 0 \\ 0 & \frac{2}{3} & \frac{1}{3}\mathrm{e}^{\gamma 3} \end{bmatrix}$ is greater than $\frac{1}{2}[\exp(\gamma 2) + 1]$ (observe that $\frac{2}{3} + \frac{1}{3}\mathrm{e}^{\gamma 3} > \frac{1}{2} + \frac{1}{2}\mathrm{e}^{\gamma 2}$ for each $\gamma > 0$).

Conclusions

In this note basic facts concerning decision under uncertainty along with typical utility functions are summarized. Special attention is focused on properties of the expected utility and the corresponding certainty equivalent if the stream of obtained rewards is governed by Markov dependence and evaluated by exponential utility functions.

Acknowledgement. This research was supported by the Czech Science Foundation under Grants P402/11/0150, P402/10/1610 and P402/10/0956.

References

[1] J. von Neumann and O. Morgenstern: Theory of Games and Economic Behaviour. Third Edition, Princeton Univ. Press, Princeton, NJ 1953.

[2] D. P. Bertsekas: Dynamic Programming and Optimal Control, Volume 1. Third Edition. Athena Scientific, Belmont, Mass. 2007.

[3] J. L. Corner and P. D. Corner: Characterization of decision in decision analysis practice. J. Oper. Res. Soc. *46* (1995), 304–314.

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

[5] R. A. Howard and J. Matheson: Risk-sensitive Markov decision processes. Manag. Sci. *23* (1972), 356–369.

[6] C. W. Kirkwood: Approximating risk aversion in decision analysis applications. Decision Analysis *1* (2004), 51–67.

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

[8] K. Sladký and M. Sitař: Risk sensitive and mean variance optimality in Markov decision processes. In: Proceedings of the 26th International Conference Mathematical Methods in Economics 2008 (P. Řehořová, K. Maršíková, eds.). Faculty of Economics, Technical University of Liberec, Liberec 2008, pp. 452–461.

[9] K. Sladký: Growth rates and average optimality in risk-sensitive Markov decision chains. Kybernetika *44* (2008), 206–217.

[10] A. Takayama: Analytical Methods in Economics. Harvester Wheatsheaf, Hertfordshire 1994.