McGill University
Department of Electrical & Computer Engineering

FINAL PROJECT
ECSE 506
WINTER 2020

Review of
"Two Characterizations of Optimality in Dynamic Programming"
(I. Kratzas, W.D. Sudderth) (2010)

ALPER OKER

ID: 260899205

a) Introduction

Dubins and Savage [1] showed that a strategy is optimal if and only if the strategy is both "thrifty" and "equalizing". Dubins and Savage [1] showed this result for the special set-up of a gambling problem. Since dynamic programming is very useful in economics, a different characterization which is necessary and sufficient for optimality for a specific class of economic problems with certain assumptions on the problem is adapted for economic models. For this set-up of economic problems it has been showed by Stokey and Lucas [2] that in these models with prior assumptions in order to make the model compatible with the real economic set-up, the plan is optimal if and only if the plan for the "law of motion" satisfies "Euler equation" and "Transversality Condition". So until this paper by I. Kratzas and W.D. Sudderth [0].

Dubins and Savage [1] showed the characterization

- A plan is optimal if and only if it is "thrifty" and "equalizing"

  and Blackwell [3], Hordijk [4], Rieder [5] and many others interested in dynamic programming for economical applications showed the characterization

- A plan is optimal if and only if it satisfies the "Euler equation" and the "Transversality Condition".

Objective of this paper [0] is to explain these characterizations and to show the relationship between these characterizations with the aim of making dynamic programmers working in economics more familiar with the "thrifty" and "equalizing" properties of policies.

b) Section 2: Proofs and Their Results:

Definition 1r: A discrete time supermartingale is sequence $x_1, x_2, x_3, \ldots$ of integrable random variables satisfying

$$\mathbb{E}[x_{n+1}|x_1, \ldots, x_n] \leq x_n \quad \forall_n \quad n \geq 1 \quad [6]$$

Theorem 1r (Doob's martingale convergence theorem): Let $x_1, x_2, x_3, \ldots$ be a supermartingale. Suppose that the supermartingale is bounded such that

$$\sup_{t \in N} \mathbb{E}[x_t^-] < \infty$$

where x_t^- is the negative part of x_t, defined by $x_t^- = -\min(x_t, 0)$. Then the sequence converges almost surely to a random variable with finite expectation [6].

Lemma 1p: For every plan Π and initial state s, the adapted sequences $\{M_n, F_n\}_{n\geq 1}$ and $\{\beta^{n-1}V(s_n), F_n\}_{n\geq 1}$ are non-negative supermartingales under $P^{\Pi,s}$.

Proof: Set $Q_0 = 0$, $Q_n = \beta^{n-1}r(s_n, a_n) + \sum_{k=1}^{n-1}\beta^{k-1}r(s_k, a_k)$ where $\sum_{k=1}^{n-1}\beta^{k-1}r(s_k, a_k) = Q_{n-1}$ then $Q_n = Q_{n-1} + \beta^{n-1}r(s_n, a_n)$

$$M_{n+1} = Q_n + \beta^n V(s_{n+1}) = Q_{n-1} + \beta^{n-1}r(s_n, a_n) + \beta^n V(s_{n+1}) = Q_{n-1} + \beta^{n-1}(r(s_n, a_n) + \beta V(s_{n+1}))$$

Property 1r: If X is H-measurable then $\mathbb{E}(X|H) = X$. Let F_n be the σ-field generated by the history $h_n = (s_1, a_1, \ldots, s_{n-1}, a_{n-1}, s_n)$.

Fact 1r: For a given plan Π if the history $h_n = (s_1, a_1, \ldots, s_{n-1}, a_{n-1}, s_n)$ is known then a_n is also known such that

σ-field generated by $(s_1, a_1, \ldots, s_{n-1}, a_{n-1}, s_n)$ = σ-field generated by $(s_1, a_1, \ldots, s_{n-1}, a_{n-1}, s_n, a_n)$

Now find $\mathbb{E}^{\Pi,s}[M_{n+1}|F_n]$. Q_{n-1} is a function of the variables $(s_1, a_1, \ldots, s_{n-1}, a_{n-1})$. $r(s_n, a_n)$ is a function of the variables (s_n, a_n). F_n = σ-field generated by $(s_1, a_1, \ldots, s_{n-1}, a_{n-1}, s_n)$ = σ-field generated by $(s_1, a_1, \ldots, s_{n-1}, a_{n-1}, s_n, a_n)$, by using Fact 1r.

$$\mathbb{E}^{\Pi,s}[M_{n+1}|F_n] = \mathbb{E}^{\Pi,s}[Q_{n-1}|F_n] + \beta^{n-1}(\mathbb{E}^{\Pi,s}[r(s_n, a_n)|F_n] + \beta\mathbb{E}^{\Pi,s}[V(s_{n+1})|F_n])$$

$\boxed{Q_{n-1} \text{ is } F_n\text{-measurable because } F_n \supset Q_{n-1}.\ r(s_n, a_n) \text{ is } F_n\text{-measurable as clearly seen from Fact 1r.}}$

$V(s_{n+1})$ is a function of s_{n+1} but F_n does not contain s_{n+1}. But F_n contains a_{n-1} and s_n also with given Π, F_n contains a_n. Moreover, we are given probability distribution of the next state i.e. $s_{n+1} \sim q(\cdot|s_n, a_n)$. As a result of these we get,

$$\begin{aligned}\mathbb{E}^{\Pi,s}[M_{n+1}|F_n] &= Q_{n-1} + \beta^{n-1}\left(r(s_n, a_n) + \beta \int\limits_{s_{n+1}\in S} V(s_{n+1})q(s_{n+1}|s_n, a_n)\, ds_{n+1}\right) \\ &= Q_{n-1} + \beta^{n-1}(T_{a_n}V)(s_n)\end{aligned}$$

From (2.4) $M_n = Q_{n-1} + \beta^{n-1}V(s_n)$, also we have $V(s_n) = \sup\limits_{a\in A(s_n)} (T_aV)(s_n) \geq (T_{a_n}V)(s_n)$. Then we get,

$$\mathbb{E}^{\Pi,s}[M_{n+1}|F_n] = Q_{n-1} + \beta^{n-1}(T_{a_n}V)(s_n) \leq Q_{n-1} + \beta^{n-1}V(s_n) = M_n$$

Thus, $\{M_n, F_n\}_{n\geq1}$ is a $P^{\Pi,s}$ supermartingale from Definition 1r.

Now look at the sequence $\{\beta^{n-1}V(s_n), F_n\}_{n\geq1}$.

$$\beta^{n-1}V(s_n) = M_n - Q_{n-1}$$

$\beta^nV(s_{n+1}) = M_{n+1} - Q_n$, we need to show that $\mathbb{E}^{\Pi,s}[\beta^nV(s_{n+1})|F_n] \leq \beta^{n-1}V(s_n)$ which is equivalent to showing $\mathbb{E}^{\Pi,s}[M_{n+1} - Q_n|F_n] \leq M_n - Q_{n-1}$ we need to show this.

For a given plan Π, Q_n is F_n-measurable. Then we get $\mathbb{E}^{\Pi,s}[M_{n+1} - Q_n|F_n] = \mathbb{E}^{\Pi,s}[M_{n+1}|F_n] - Q_n$.

In previous part, we showed that $\{M_n, F_N\}_{n\geq1}$ is a $P^{\Pi,s}$ supermartingale sequence. Also in the paper [0], it is assumed that daily reward function is non-negative. Then we have $Q_n \geq Q_{n-1}$. So we have $\mathbb{E}^{\Pi,s}[M_{n+1}|F_n] \leq M_n$ and $Q_n \geq Q_{n-1}$. Then $\mathbb{E}^{\Pi,s}[M_{n+1}|F_n] + Q_{n-1} \leq M_n + Q_n \Rightarrow \mathbb{E}^{\Pi,s}[M_{n+1}|F_n] - Q_n \leq M_n - Q_{n-1}$. Proof is done. $\{\beta^{n-1}V(s_n), F_n\}_{n\geq1}$ is also a $P^{\Pi,s}$ supermartingale.

Property 2r: $\mathbb{E}(\mathbb{E}(X|H)) = \mathbb{E}(X)$. This property is called "smoothing" property or "law of total expectations". [6] where H is a σ-field.

From Lemma 1p we have $\mathbb{E}^{\Pi,s}[M_{n+1}|F_n] \leq M_n$. Now take the the expected value of both sides and use smoothing property, we get

$$\mathbb{E}^{\Pi,s}[\mathbb{E}^{\Pi,s}[M_{n+1}|F_n]] \leq \mathbb{E}^{\Pi,s}[M_n] \Rightarrow \boxed{\mathbb{E}^{\Pi,s}[M_{n+1}] \leq \mathbb{E}^{\Pi,s}[M_n]}$$

Same goes for $\{\beta^{n-1}V(s_n)\}_{n\geq1}$. So the sequences $\{M_n\}_{n\geq1}$ and $\{\beta^{n-1}V(s_n)\}$ are non-increasing in expectation by the virtue of smoothing property and converge almost surely by the Theorem 1r (Doob's martingale convergence theorem).

Now I will prove the chain of equalities and inequalities in equation (2.6).

Fix an initial $s = s_1$. Then $M_1 = V(s_1)$. Here it is useful to note that M_1 is not a random variable for a fixed $s = s_1$. M_1 is a constant value. I will use the following property: "Expectation of a constant is the constant itself i.e. $\mathbb{E}(a) = a$ if 'a' is constant". Then,

$$V(s_1) = M_1 = \mathbb{E}^{\Pi,s}(M_1) \Rightarrow V(s) = \mathbb{E}^{\Pi,s}(M_1)$$

Suppose the initial state is s. So we proved first equality. Since $s_1 = s$ is arbitrary I will write s. Since the sequence $\{M_n\}_{n\geq1}$ is non-increasing in expectation by smoothing property we get,

$$\begin{aligned}\mathbb{E}^{\Pi,s}(M_1) \geq \lim_{n\to\infty} \mathbb{E}^{\Pi,s}(M_{n+1}) = \Lambda^\Pi(s)& \\ &= \lim_{n\to\infty}\{\mathbb{E}^{\Pi,s}(Q_n) + \beta^n\mathbb{E}^{\Pi,s}[V(s_{n+1})]\} \\ &= R^\Pi(s) + \lim_{n\to\infty}\{\beta^n\mathbb{E}^{\Pi,s}[V(s_{n+1})]\} \geq R^\Pi(s)\end{aligned}$$

Now I will explain why the last inequality holds. The equalities in between which I did not show are direct results of definitions. From assumptions of this paper, daily reward function $r(\cdot,\cdot)$ is non-negative.

$$V(s) = \sup_{\Pi} \mathbb{E}^{\Pi,s}\left(\sum_{n=1}^{\infty} \beta^{n-1} \cdot r(s_n, a_n)\right)$$

Since $r(\cdot,\cdot)$ is non-negative, this value is also non-negative. So $V(s) \geq 0$ for $\forall\, s$. As a result,

$$\lim_{n\to\infty} \{\beta^n \mathbb{E}^{\Pi,s}[V(s_{n+1})]\} \geq 0.$$

We proved last inequality of (2.6).

Definition 1p: A given plan Π is called: thrifty at $s \in S$, if $V(s) = \Lambda^{\Pi}(s)$; it is called equalizing at $s \in S$ if $\Lambda^{\Pi}(s) = R^{\Pi}(s)$.

Theorem 1p: A plan Π is optimal at $s \in S$ if and only if Π is both thrifty and equalizing at s.

Proof: First prove,

- If a plan Π is optimal at $s \in S$ then Π is both thrifty and equalizing at s. Suppose Π is optimal. Then we have
$$R^{\Pi}(s) = V(s)$$
Also from (2.6) we have
$$\boxed{V(s) \geq \Lambda^{\Pi}(s) \geq R^{\Pi}(s)}$$
So the only way that both of these expressions are satisfied is
$$\boxed{V(s) = \Lambda^{\Pi}(s) = R^{\Pi}(s)}$$
From Definition 1p, we see that if Π is optimal then Π is both thrifty and equalizing at s.

Secondly prove,

- If a plan Π is both thrifty and equalizing at s then Π is optimal at s. Suppose Π is both thrifty and equalizing, by thrifty property we have $V(s) = \Lambda^{\Pi}(s)$, by equalizing property we have $\Lambda^{\Pi}(s) = R^{\Pi}(s)$.
So we get $R^{\Pi}(s) = V(s)$, then Π is optimal at $s \in S$. Proof of Theorem 1p is complete.

Theorem 2p: For a given plan Π and initial state $s \in S$, the following are equivalent:

(a) the plan Π is thrifty at s

(b) the sequence $\{M_n, F_n\}_{n\geq 1}$ is a martingale under $P^{\Pi,s}$

(c) for all $n \geq 1$, we have $P^{\Pi,s}(a_n$ conserves $V(\cdot)$ at $s_n) = 1$

Proof: To prove Theorem 2p, we need to prove the following, $a \Rightarrow b$, $b \Rightarrow c$, $c \Rightarrow a$. Start by proving $a \Rightarrow b$. Suppose the plan Π is thrifty at s, then $V(s) = \Lambda^{\Pi}(s)$. Also from (2.6) and non-increasing property of the sequence $\{M_n\}_{n\geq 1}$ in expectation, we get,

$$\begin{aligned} V(s) &= \mathbb{E}^{\Pi,s}(M_1) = \mathbb{E}^{\Pi,s}(M_2) = \mathbb{E}^{\Pi,s}(M_3) = \cdots = \mathbb{E}^{\Pi,s}(M_k) = \dots \\ &= \lim_{n\to\infty} \mathbb{E}^{\Pi,s}(M_n) = \Lambda^{\Pi}(s) \end{aligned}$$

Now check if $\{M_n, F_n\}_{n\geq 1}$ is supermartingale, submartingale or martingale by using above chain of equality, first suppose it is supermartingale then $\mathbb{E}^{\Pi,s}[M_{n+1}|F_n] \leq M_n$ use smoothing $\mathbb{E}^{\Pi,s}[M_{n+1}] \leq \mathbb{E}^{\Pi,s}[M_n]$ which contradicts the above chain of equation. So $\{M_n, F_n\}_{n\geq 1}$ is not supermartingale.

With same kind of argument, $\{M_n, F_n\}_{n\geq 1}$ is not submartingale. So this chain of equalities

$$\mathbb{E}^{\Pi,s}(M_1) = \mathbb{E}^{\Pi,s}(M_2) = \cdots = \mathbb{E}^{\Pi,s}(M_k) = \cdots = \lim_{n\to\infty} \mathbb{E}^{\Pi,s}(M_n)$$

implies that $\{M_n, F_n\}_{n\geq 1}$ is martingale under $P^{\Pi,s}$. We proved $a \Rightarrow b$. Now prove $b \Rightarrow c$. Suppose the sequence $\{M_n, F_n\}_{n\geq 1}$ is a martingale under $P^{\pi,s}$, then

$$\mathbb{E}^{\Pi,s}[M_{n+1}|F_n] = M_n \quad \text{for all} \quad n \geq 1$$

Using the chain of equality and inequalities in (2.5) we get

$$Q_{n-1} + \beta^{n-1}(T_{a_n}V)(s_n) = Q_{n-1} + \beta^{n-1}V(s_n)$$

from this we get

$$\boxed{(T_{a_n}V)(s_n) = V(s_n)} \quad \text{for all} \quad n \geq 1$$

So this shows that a_n conserves $V(\cdot)$ at s_n for all $n \geq 1$. Thus we proved $b \Rightarrow c$.

Now prove $c \Rightarrow a$. Suppose a_n conserves $V(\cdot)$ at s_n for all $n \geq 1$. Then

$$(T_{a_n}V)(s_n) = V(s_n) \quad \text{for all} \quad n \geq 1.$$

This equality implies that $\mathbb{E}^{\Pi,s}[M_{n+1}|F_n] = M_n$ and this implies by smoothing that $\mathbb{E}^{\Pi,s}[M_{n+1}] = \mathbb{E}^{\Pi,s}[M_n]$ for all $n \geq 1$. Plug this equality into (2.6) and we get $V(s) = \Lambda^{\Pi}(s)$ which implies the plan Π is thrifty at s. So $a \Rightarrow b \Rightarrow c \Rightarrow a$ holds. Then statements a, b and c are equivalent.

Theorem 3p: A given plan Π is equalizing at $s \in S$ if and only if we have $\lim_{n\to\infty}\left(\beta^n \mathbb{E}^{\Pi,s}[V(s_{n+1})]\right) = 0$.

Proof: First prove, if Π is equalizing then $\lim_{n\to\infty}\left(\beta^n \mathbb{E}^{\Pi,s}[V(s_{n+1})]\right) = 0$. Suppose Π is equalizing. This means that $\Lambda^{\Pi}(s) = R^{\Pi}(s)$, $\Lambda^{\Pi}(s) = R^{\Pi}(s) + \lim_{n\to\infty}\{\beta^n \mathbb{E}^{\Pi,s}[V(s_{n+1})]\} = R^{\Pi}(s)$. Then $\lim_{n\to\infty}\{\beta^n \mathbb{E}^{\Pi,s}[V(s_{n+1})]\} = 0$. So "$\Rightarrow$" is proved. Now prove, if $\lim_{n\to\infty}\left(\beta^n \mathbb{E}^{\Pi,s}[V(s_{n+1})]\right) = 0$ then Π is equalizing. Suppose $\lim_{n\to\infty}\left(\beta^n \mathbb{E}^{\Pi,s}[V(s_{n+1})]\right) = 0$. Then $\Lambda^{\Pi}(s) = R^{\Pi}(s)$. So "$\Leftarrow$" is proved. Proof is done.

Now assume that $0 \leq r(s,a) \leq K < \infty$ i.e. daily reward function is bounded. Suppose $r(s_n, a_n) = K$ for all $n \geq 1$ in order to find the maximum reward that can be achieved. Then we get

$$\begin{aligned}\lim_{n\to\infty}(K + \beta K + \cdots + \beta^n K) &= \lim_{n\to\infty}(K(\beta^n + \beta^{n-1} + \cdots + 1)) \\ &= \lim_{n\to\infty}\left(\frac{K(\beta^{n+1}-1)}{(\beta-1)}\right) = \frac{-K}{\beta-1} = \frac{K}{1-\beta}.\end{aligned}$$

So $0 \leq V(s) \leq K/(1-\beta)$ for all $s \in S$ and every plan is equalizing for the special case of bounded daily reward case.

Another way to interpret "thrifty" property is by looking at Theorem 2p (c). Theorem 2p (c) states that a plan Π is thrifty if, with probability one i.e. $P^{\Pi,s} = 1$, it makes no "immediate, irremediable mistakes" along any history.

Another way to interpret equalizing property is by looking at Theorem 3p. Theorem 3p states that a plan Π is equalizing if "eventually it is certain to force the system into states where little further income can be anticipated".

Theorem 4p: Let $V_0(\cdot)$ be identically zero. Then for all $s \in S$ and $n = 1, 2, \ldots$

(a) $V_{n+1}(s) = \sup_{a\in A(s)} (T_a V_n)(s)$

(b) $V(s) = \lim_{n\to\infty} V_n(s)$

Proof: Proof of this theorem follows from ESCE 506 lecture notes on Infinite Horizon Discounted MDP. By Theorem 1, Proposition 2 in the notes and also from Banach fixed point theorem this can be proved.

Pros and Cons of Equalizing and Thrifty Characterization:

- Holds in great generality.
- Often useful for proving general properties: For example, if every $A(s)$ (set of actions available at s) is finite and r(daily reward function) is bounded, then there is an optimal stationary plan.
- Typically hard to apply to specific cases: We usually need to know the optimal reward function V in order to check the conditions.

c) Section 3: Proofs and Results:

Lemma 2p: The value function $V(y,z)$ is concave in y; hence so is $\Psi(x,y,z)$. The function $\Psi(x,y,z)$ is strictly concave in y if $F(x,y,z)$ is.

Proof: In assumptions of the problem, we are given that $F(x,y,z)$ is concave in the pair (x,y) for every given $z \in Z$. Hence we assume $F(x,y,z)$ is concave both in its first argument x and its second argument y. Also from Theroem 4p, we assume that $V_0(\cdot,\cdot)$ is identically zero. Proof will be done by forward induction.

$$V_1(x,z) = \sup_{y\in\Gamma(x,z)} F(x,y,z) \quad \text{since} \quad V_0(\cdot,\cdot) = 0. \text{ Then}$$

$$V_1(x,z) = \sup\Big\{F(x,y_1,z), F(x,y_2,z), F(x,y_3,z), \ldots, F(x,y_{k-1},z), F(x,y_k,z)\Big\}$$

So it can be clearly seen that $V_1(x,z)$ will be equal to one of the arguments in the $\sup\{\cdot,\cdot,\ldots,\cdot,\cdot\}$ operator, where $y_1, y_2, \ldots, y_k$ are fixed constant values. But we know that all of these arguments are concave in their first arguments x. So $V_1(x,z)$ is concave in its first argument. Then we have

$$\left.\begin{array}{ll} V_0(x,z) = 0 & \text{trivially concave in x} \\ V_1(x,z) & \text{concave in x} \end{array}\right\} \text{Basis for induction}$$

So now we have basis for induction. Now suppose that $V_n(x,z)$ is concave in its first argument x. Then

$$V_{n+1}(x,z) = \sup_{y\in\Gamma(x,z)} \left(F(x,y,z) + \beta \int_Z V_n(y,t)q(dt|z)\right)$$

First look at the term $\beta \int_Z V_n(y,t)q(dt|z)$. This term is a function of y. Call this term $\Phi(y)$. Then we get

$$V_{n+1}(x,z) = \sup_{y\in\Gamma(x,z)} \Big(F(x,y,z) + \Phi(y)\Big)$$
$$V_{n+1}(x,z) = \sup\left\{\Big(F(x,y_1,z) + \Phi(y_1)\Big), \Big(F(x,y_2,z) + \Phi(y_2)\Big), \ldots, \Big(F(x,y_k,z) + \Phi(y_k)\Big)\right\}$$

So it can be seen that $V_{n+1}(x,z)$ will be equal to one of the arguments in supremum. Since all $y_1, y_2, \ldots, y_k$ are constant values, all arguments in the supremum are concave in their first arguments x. Then $V_{n+1}(x,z)$ is also concave in its first argument x. Since it is concave in its first argument then $V_{n+1}(y,z)$ is also concave in its first argument y. From Theorem 4p (b), the function $V(y,z)$ is the pointwise limit of concave function i.e. $\lim_{n\to\infty} V_n(y,z) = V(y,z)$. Therefore $V(y,z)$ is also concave in its first argument y. Now look at $\Psi(x,y,z)$.

$$\Psi(x,y,z) = F(x,y,z) + \beta \int_Z V(y,t)q(dt|z)$$

$\beta \int_Z V(y,t)q(dt|z)$ is a weighted sum of concave functions, $V(y,t)$, which is concave in its first arguments y. Also all weights are positive because $q(dt|z)$ is a probability distribution. Since all weights are positive and this is a weighted sum of concave functions $V(y,t)$, which is concave in its first argument y, then

$$\boxed{\beta \int_Z V(y,t)q(dt|z) \quad \text{is concave in } y.}$$

Also from initial assumption we have that $F(x, y, z)$ is concave in its second argument y. The summation of two concave functions is concave then $\Psi(x, y, z)$ is concave in its second argument y. Proof is done.

Lemma 3p:

(i) The value function $V(x, z)$ is non-decreasing in x.

(ii) If the partial derivatives $D_x V(x, t)$ exist for $q(t|z)$ for almost all $t \in Z$, then whenever both its sides are well-defined, the following equality holds:

$$D_x \int_Z V(x,t)q(dt|z) = \int_Z D_x V(x,t)q(dt|z)$$

Proof:

(i) Let $x \le x'$. Consider a plan Π for the player who begins at state (x, z). Suppose this player who started at (x, z) choose the initial action y. Then this (x, z) player gets initial daily reward $F(x, y, z)$ and moves to state (y, t) where t is distributed as $t \sim q(\cdot|z)$. Now think of the player who started at (x', z). Suppose this (x', z) player also chooses y as its initial action. Then this (x', z) player gets initial daily reward $F(x', y, z)$ and moves to state (y, t) where t is distributed as $t \sim q(\cdot|z)$.
Now suppose that after choosing their initial actions as y (both of the players) they choose the same actions afterwards. In other words after first common action y, both players chooses the same actions all times (from $n = 2$ to $n = \infty$). But after first action, both players goes to the same state. So after first action everything becomes identical for both players when they choose the same actions. From initial assumptions, we have that $F(x, y, z)$ is non-decreasing in x. So at first step

$$\left.\begin{array}{lll} (x,z) & \text{player gets} & \text{F(x,y,z)} \\ (x',z) & \text{player gets} & \text{F(x}',\text{y,z)} \end{array}\right\} \text{Since } x' \ge x \text{ then } F(x',y,z) \ge F(x,y,z)$$

So at first step (x', z) player gets more reward than (x, z) player but afterwards they get the same rewards. So we showed that whatever the plan Π of (x, z) player is, (x', z) player can achieve at least, as much total discounted reward as (x, z) player. Suppose the plan Π of (x, z) player is optimal i.e. total discounted reward of (x, z) player is $V(x, z)$. But we showed that (x', z) player can achieve at least, as much total discounted reward as (x, z) player. Then we proved that

$$V(x',z) \ge V(x,z) \quad \text{for } x' \ge x$$

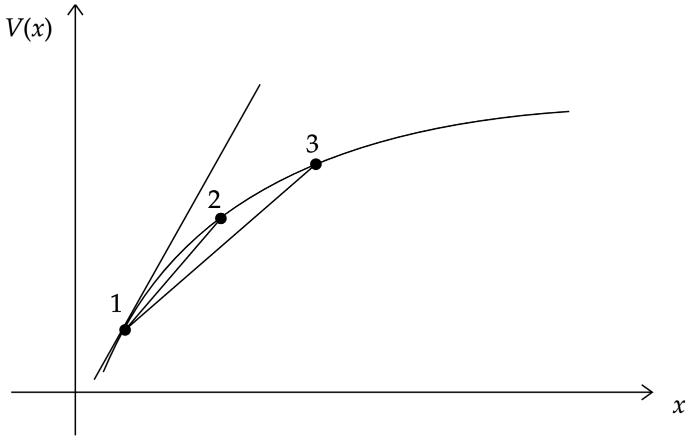
Since x' was arbitrary, this result holds for all $x' \ge x$.

Proof:

(ii) For $\varepsilon > 0$ consider the quotients

$$\Big(V(x_1,\dots,x_i+\varepsilon,\dots,x_L,z) - V(x_1,\dots,x_i,\dots,x_L,z)\Big)/\varepsilon$$

By part (i) of Lemma 3p, these are non-negative. By concavity of $V(\cdot, z)$, they are non-decreasing as $\varepsilon \downarrow 0$ [7]. I want to show intuitively why this is the case on a simple 2-dimensional example. This quotient simply represents $D_x V(x, z)$ when $\varepsilon \to 0$. So if this quotient is non-decreasing as $\varepsilon \to 0$ by concavity then we can conclude that this quotient sequence as $\varepsilon \to 0$ is monotone because it is non-decreasing. Also from assumptions we have that $F(x, y, z)$ is continuously differentiable then $V(x, z)$ is also continuously differentiable. Then we have that quotients as $\varepsilon \to 0$ are bounded. Since quotients are monotone and bounded we have monotone convergence.



$1 = V(x)$, $2 = V(x + \varepsilon_1)$, $3 = V(x + \varepsilon_2)$ where $\varepsilon_2 > \varepsilon_1$. As seen as $\varepsilon \to 0$, the quotient becomes the slope of the tangent line to the curve at 1 i.e. the derivative at 1. When $\varepsilon = \varepsilon_2$ the quotient becomes the slope of the line which connects 1 and 3. When $\varepsilon = \varepsilon_1$ the quotient becomes the slope of the line which connects 1 and 2. AS can be seen in the graph, as ε goes from $\varepsilon_2 \to \varepsilon_1 \to \varepsilon \to 0$ the slope increases. So for a concave function quotient increases as $\varepsilon \to 0$.

From monotone convergence theorem, we know that if a sequence is monotone convergent then we have the following

$$\lim_{\varepsilon \to 0} \int (\cdots) = \int \lim_{\varepsilon \to 0} (\cdots)$$

So D_x operator which represents derivative (gradient for a vector of variables) can commute between both sides of the integral operator. So proof of (ii) is done.

Theorem 5p: Suppose that Π is an interior plan at $s = (x, z)$. Then Π is thrifty at s if and only if the following hold with probability one under $P^{\Pi,s}$:

(a) The envelope equation

$$D_x V(x_n, z_n) = D_x F(x_n, y_n, z_n), \quad \forall n = 1, 2, \ldots;$$

(b) The Euler equation

$$D_y F(x_n, y_n, z_n) + \beta \int\limits_Z D_x F(y_n, y_{n+1}, t) q(dt|z_n) = 0 \quad \forall n = 1, 2, \ldots$$

Proof: First prove the forward ($\Rightarrow$) statement. So first assume Π is thrifty. By Theorem 2p, the actions y_n conserve $V(\cdot, \cdot)$ at s_n for all $n \in N$, on an event of probability one. Hence, y_n maximizes $\Psi(x_n, \cdot, z_n)$ over $\Gamma(x_n, z_n)$, for all $n \in N$. In order to prove this, first we need to show that $V(\cdot, z_n)$ is differentiable at x_n. By the interiority and continuity assumptions, there exits for every $n \in N$, an open neighborhood O_n of x_n such that $y_n \in \Gamma(x, z_n)$ holds for $\forall x \in O_n$. Therfore the function

$$W(x) = F(x, y_n, z_n) + \beta \int\limits_Z V(y_n, t) q(dt|z_n)$$

is concave, continuously differentiable and satisfies $W(x) \le V(x, z_n)$, $\forall x \in O_n$ with equality for $x = x_n$. Then we get

$$V(x, z_n) - V(x_n, z_n) \ge W(x) - W(x_n), \quad \forall x \in O_n$$

Property 3r: The concepts of subderivative and subdifferential can be generalized to functions of several variables. If $f : U \to \mathbb{R}$ is a real-valued concave function defined on a convex set in the Euclidean space $\mathbb{R}^n$, a vector v in that space is called a subgradient at a point x_0 in U if for any x in U one has

$$V \cdot (x - x_0) \ge f(x) - f(x_0)$$

where the dot denotes the dot product. [6]

Property 4r: A concave function $f : I \to \mathbb{R}$ is differentiable at x_0 if and only if the subgradient is made up of only one point, which is the derivative at x_0 [6]. Now, any subgradient $p \in \mathbb{R}^\ell$ of the function $V(\cdot, z_n)$ at x_n must satisfy

$$p \cdot (x - x_n) \geq V(x, z_n) - V(x_n, z_n) \geq W(x) - W(x_n) \quad \forall x \in O_n$$

where the dot "$\cdot$" represents inner product in $\mathbb{R}^\ell$. By interiority and continuity assumptions, we have that $W(\cdot)$ is differentiable at x_n, so by Property 4r, there is only one such subgradient. This means that the value of p is unique i.e. p has a single value which is the derivative of $W(\cdot)$ at x_n. Since p has a unique value this implies by Property 4r that $V(\cdot, z_n)$ is also differentiable at x_n. Since the p is the value of the derivative at x_n for both $V(\cdot, z_n)$ and $W(\cdot, z_n)$, then they have the same derivative at x_n. Since $V(\cdot, z_n)$ is differentiable at $x = x_n$, it is legitimate to wirte

$$D_x V(x_n, z_n) = D_x W(x_n)$$

So we proved that $V(\cdot, z_n)$ is differentiable at x_n. Now think of the equality,

$$\Psi(x, y, z) = F(x, y, z) + \beta \int_Z V(y, t) q(dt|z)$$

Since $\Psi(x, y, z)$ is maximized by $y = y_n$ we get

$$\Psi(x, y_n, z) = V(x, z), \quad \text{then}$$

$$V(x, z) = F(x, y_n, z) + \beta \int_Z V(y_n, t) q(dt|z)$$

$\beta \int_Z V(y_n, t) q(dt|z)$ is a constant, call it c =constant then,

$$V(x, z) = F(x, y_n, z) + c$$

By assumption we have that $F(x, y, z)$ is continuously differentiable in the interior of $X \times X$ for $\forall z \in Z$. Also in previous part we showed that $V(\cdot, z_n)$ is differentiable at $x = x_n$. Then by taking the gradient of both sides of the above equation at $x = x_n$ we get

$$D_x V(x_n, z) = D_x F(x_n, y_n, z) + D_x c$$

Since c is a constant $D_x c = 0$. Then we get

$$D_x V(x_n, z_n) = D_x F(x_n, y_n, z_n) \quad \forall n = 1, 2, \ldots;$$

So we proved that, if Π is thrifty then the envelope equation holds. Now assume Π is thrifty and try to prove (b) the Euler equation. Since Π is thrifty, we have that y_n maximizes $\Psi(x, y, z)$ over y. Then the derivative with respect to y at $y = y_n$ gives

$$D_y \Psi(x_n, y_n, z_n) = 0, \quad \text{For simplicity take } x = x_n,\ z = z_n$$

$$\Psi(x, y, z) = F(x, y, z) + \beta \int_Z V(y, t) q(dt|z)$$

Take the derivative of both sides with respect to y (gradient if y is a vector). Then we get

$$D_y \Psi(x, y, z) = D_y F(x, y, z) + \beta D_x \int_Z V(y, t) q(dt|z)$$

By using Lemma 3p (ii) we get

$$D_y \Psi(x, y, z) = D_y F(x, y, z) + \beta \int_Z D_x V(y, t) q(dt|z)$$

Evaluating this expression at $y = y_n$ we get

$$D_y F(x, y_n, z) + \beta \int_Z D_x V(y_n, t) q(dt|z) = 0$$

By using Envelope equality, we get

$$D_yF(x, y_n, z) + \beta \int\limits_Z D_xF(y_n, y_{n+1}, t)q(dt|z_n) = 0$$

Here z represents $z = z_n$, x represents $x = x_n$. Then we have

$$D_yF(x_n, y_n, z_n) + \beta \int\limits_Z D_xF(y_n, y_{n+1}, t)q(dt|z_n) = 0, \quad \forall n \geq 1$$

So we proved that if Π is thrifty the Euler equation holds. Proof of forward statement is complete.

Since the σ-field F_n is generated by history h_n which contains z_n, the Euler equation can be written equivalently as

$$D_yF(x_n, y_n, z_n) + \beta \mathbb{E}^{\Pi,s}[D_xF(y_n, y_{n+1}, z_{n+1})|F_n] = 0$$

Now prove the converse statement. Assume that Envelope and Euler equations hold. We need to show that, with $P^{\Pi,s}$-probability one, y_n maximizes on the set $\Gamma(x_n, z_n)$ the concave function $\Psi(x_n, \cdot, z_n)$ for each $n \in N$. From Euler and Envelope equalities and Lemma 3p (ii) we have (recalling $y_n \equiv x_{n+1}$)

$$\begin{aligned} D_y\Psi(x_n, y_n, z_n) &= D_yF(x_n, y_n, z_n) + \beta \int\limits_Z D_xV(y_n, t)q(dt|z_n) \\ &= D_yF(x_n, y_n, z_n) + \beta \int\limits_Z D_xF(y_n, y_{n+1}, t)q(dt|z_n) = 0 \end{aligned}$$

In this chain equations we used Euler and Envelope equalities, also Lemma 3p (ii).

Since $\Psi(x, y, z)$ is concave in its second argument, y, if gradient with respect to y is equal to zero at some point y_n then this point y_n is a maximizer. So, gradient being equal to zero is necessary and sufficient condition for y_n to be a maximizer for a concave function. Also for concave function all local maximizers are also global maximizers. Then from previous part

$$D_y\Psi(x_n, y_n, z_n) = 0$$

shows us that y_n maximizes the concave function $\Psi(x_n, \cdot, z_n)$ on the set $\Gamma(x_n, z_n)$ with probability 1 for $\forall n \in N$. Then the converse proof is done, Π is thrifty. Since we proved both forward and converse statements, we proved "if and only if" statement.

Theorem 6p: Suppose the plan Π is optimal and interior at $s = (x, z)$ and that the reward function satisfies the requirement

$$x \cdot D_xF(x, y, z) \geq 0$$

for all interior states (x, z) and interior actions $y \in \Gamma(x, z)$. Then Π satisfies
(c) the Transversality condition

$$\lim_{n\to\infty} \left(\beta^n \mathbb{E}^{\Pi,s}[x_nD_xF(x_n, y_n, z_n)] \right) = 0$$

Proof:
Property 5r: If f is concave and differentiable, then it is bounded above by its first-order Taylor approximation $f(y) \leq f(x) + f'(x)[y - x]$ which is equivalent to $f(x) - f(y) \geq [x - y]f'(x)$ [6]. $V(x_n, z_n) \geq V(x_n, z_n) - V(0, z_n)$. This holds because $V(\cdot, \cdot)$ is non-negative.

$$V(x_n, z_n) - V(0, z_n) \geq x_n \cdot D_xV(x_n, z_n)$$

This holds by Property 5r.

$$x_n \cdot D_xV(x_N, z_n) = x_n \cdot D_xF(x_n, y_n, z_n)$$

Since Π is optimal, then it is thrifty. Since it is thrifty then Envelope equation holds. So the above equation holds by Envelope equality. Then we get

$$V(x_n, z_n) \geq V(x_n, z_n) - V(0, z_n) \geq x_n \cdot D_xV(x_n, z_n) = x_n \cdot D_xF(x_n, y_n, z_n) \geq 0$$

Then we have

$$V(x_n, z_n) \geq x_n \cdot D_x F(x_n, y_n, z_n) \geq 0$$

Then we have

$$\beta^n \cdot V(x_n, z_n) \geq \beta^n \cdot x_n \cdot D_x F(x_n, y_n, z_n) \geq 0. \quad \text{Then}$$

$$\beta^n \mathbb{E}^{\Pi,s}[V(x_n, z_n)] \geq \beta^n \mathbb{E}^{\Pi,s}[x_n \cdot D_x F(x_n, y_n, z_n)] \geq 0. \quad \text{Then}$$

$$\lim_{n\to\infty} \Big(\beta^n \mathbb{E}^{\Pi,s}[V(x_n, z_n)]\Big) \geq \lim_{n\to\infty} \Big(\beta^n \mathbb{E}^{\Pi,s}[x_n \cdot D_x F(x_n, y_n, z_n)]\Big) \geq 0$$

Since Π is optimal then it is equalizing, so we have

$$\lim_{n\to\infty} \Big(\beta^n \mathbb{E}^{\Pi,s}[V(x_n, z_n)]\Big) = 0 \quad \text{So we have}$$

$$0 \geq \lim_{n\to\infty} \Big(\beta^n \mathbb{E}^{\Pi,s}[x_n \cdot D_x F(x_n, y_n, z_n)]\Big) \geq 0$$

This implies the transversality condition which is

$$\lim_{n\to\infty} \Big(\beta^n \mathbb{E}^{\Pi,s}[x_n \cdot D_x F(x_n, y_n, z_n)]\Big) = 0, \quad \text{proof is complete.}$$

Note: Since $F(x, y, z)$ is assumed to be non-decreasing in x then $D_x F(x, y, z) \geq 0$. So the assumption $x \cdot D_x F(x, y, z) \geq 0$ is satisfied when all states $x \in X$ lie in non-negative orhant $\mathbb{R}^\ell_+$ of $\mathbb{R}^\ell$. This is not a strict assumption for economical models since in most economic applications $x \in X$ lies in $\mathbb{R}^\ell_+$ automatically without the need for imposing a restriction on x.

Theorem 7p: Suppose the plan Π is interior at $s = (x, z)$ and $x \cdot D_x F(x, y, z) \geq 0$ holds. If Π is optimal, then it satisfies both the Euler equation with $P^{\Pi,s}$-probability one, and the transversality condition. Conversely, if these two conditions hold for Π and in addition, we have $X \subseteq \mathbb{R}^\ell_+$ then Π is optimal.

Proof: First prove forward ($\Rightarrow$) statement. Suppose Π is optimal, then Π is thrifty and by Theorem 5p it satisfies Euler equation. Since Π is optimal by Theorem 6p it satisfies transversality condition. So forward statement is proved. Now prove converse statement ($\Leftarrow$).

Suppose the plan Π satisfies both the Euler equation and transversality condition. Also suppose $X \subseteq \mathbb{R}^\ell_+$. Now consider an arbitrary plan $\widetilde{\Pi}$. So $\{Q_n\}_{n\in N}$ corresponds to the plan Π and $\{\widetilde{Q_n}\}_{n\in N}$ corresponds to the plan $\widetilde{\Pi}$. We get,

$$\beta(Q_N - \widetilde{Q_N}) = \sum_{n=1}^{N} \beta^n \big[F(x_n, y_n, z_n) - F(\tilde{x}_n, \tilde{y}_n, \tilde{z}_n)\big]$$

This equality holds by the definition of Q_N.

$$\sum_{n=1}^{N} \beta^n \big[F(x_n, y_n, z_n) - F(\tilde{x}_n, \tilde{y}_n, \tilde{z}_n)\big] \geq \sum_{n=1}^{N} \beta^n \big[(x_n - \tilde{x}_n) \cdot D_x F(x_n, y_n, z_n) + (x_{n+1} - \tilde{x}_{n+1}) \cdot D_y F(x_n, y_n, z_n)\big]$$

This inequality follows from Property 5r and $x_{n+1} = y_n$. Since $F(\cdot, \cdot, z)$ is concave we can use Property 5r.

$$\begin{aligned}
&\sum_{n=1}^{N} \beta^n \big[(x_n - \tilde{x}_n) \cdot D_x F(x_n, y_n, z_n) + (x_{n+1} - \tilde{x}_{n+1}) \cdot D_y F(x_n, y_n, z_n)\big] \\
&= \sum_{n=1}^{N} \beta^{n-1} (x_n - \tilde{x}_n) \cdot \big[D_y F(x_{n-1}, y_{n-1}, z_{n-1}) + \beta D_x F(y_{n-1}, y_n, z_n)\big] \\
&+ \beta^N (x_{N+1} - \tilde{x}_{N+1}) \cdot D_y F(x_N, y_N, z_N) \quad \forall N \in N
\end{aligned}$$

This holds by arrangement of terms in the first summation. By taking expectations we get,

$$\begin{aligned}
&\beta\big[\mathbb{E}(Q_N) - \mathbb{E}(\tilde{Q}_N)\big] \geq \beta^N \cdot \mathbb{E}\Big[(x_{N+1} - \tilde{x}_{N+1}) \cdot D_y F(x_N, y_N, z_N)\Big] \\
&+ \mathbb{E} \sum_{n=0}^{N-1} \beta^n (x_{n+1} - \tilde{x}_{n+1}) \cdot \Big[D_y F(x_n, y_n, z_n) + \beta \mathbb{E}\Big(D_x F(y_n, y_{n+1}, z_{n+1}) | F_{n+1}\Big)\Big]
\end{aligned}$$

where we have used the smoothing property of expectation. Since Π satisfies the transversality condition then we have

$$D_yF(x_n, y_n, z_n) + \beta\mathbb{E}\Big[D_xF(y_n, y_{n+1}, z_{n+1})|F_{n+1}\Big] = 0.$$

Then we have

$$\mathbb{E}(Q_N) - \mathbb{E}(\tilde{Q}_N) \geq \beta^{N-1} \cdot \mathbb{E}\Big[(x_{N+1} - \tilde{x}_{N+1}) \cdot D_yF(x_N, y_N, z_N)\Big]$$

From Euler equation we have

$$D_yF(x_n, y_n, z_n) + \beta \cdot \mathbb{E}\Big(D_xF(y_n, y_{n+1}, z_{n+1})|F_{n+1}\Big) = 0$$

Then multiply both sides with $(x_{N+1} - \tilde{x}_{N+1})$. Then we get

$$(x_{N+1} - \tilde{x}_{N+1}) \cdot D_yF(x_n, y_n, z_n) + \beta(x_{N+1} - \tilde{x}_{N+1}) \cdot \mathbb{E}\Big[D_xF(y_n, y_{n+1}, z_{n+1})|F_{n+1}\Big] = 0$$

Now take $n = N$, then we get

$$(x_{N+1} - \tilde{x}_{N+1}) \cdot D_yF(x_N, y_N, z_N) + \beta(x_{N+1} - \tilde{x}_{N+1}) \cdot \mathbb{E}\Big[D_xF(y_N, y_{N+1}, z_{N+1})|F_{N+1}\Big] = 0$$

$F_{N+1} = \sigma(x_1, y_1, z_1, \ldots, x_N, y_N, z_N, x_{N+1}, z_{N+1})$. So F_{N+1} contains the information of x_{N+1}. This means that we can commute x_{N+1} inside the expectation.

Since F_{N+1} has the information x_{N+1}, we can commute it inside the expectation then we get

$$\begin{aligned}&(x_{N+1} - \tilde{x}_{N+1}) \cdot D_yF(x_N, y_N, z_N) + \beta \cdot \mathbb{E}\Big[x_{N+1}D_xF(y_N, y_{N+1}, z_{N+1})|F_{N+1}\Big]\\ &- \beta \cdot \tilde{x}_{N+1} \cdot \mathbb{E}\Big[D_xF(y_N, y_{N+1}, z_{N+1})|F_{N+1}\Big] = 0\end{aligned}$$

Then we have

$$\begin{aligned}&(x_{N+1} - \tilde{x}_{N+1}) \cdot D_yF(x_N, y_N, z_N) + \beta \cdot \mathbb{E}\Big[x_{N+1}D_xF(y_N, y_{N+1}, z_{N+1})|F_{N+1}\Big]\\ &= \beta \cdot \tilde{x}_{N+1} \cdot \mathbb{E}\Big[D_xF(y_N, y_{N+1}, z_{N+1})|F_{N+1}\Big], \text{ multiply both sides with } \beta^{N-1}, \text{ then}\\ &\beta^{N-1}(x_{N+1} - \tilde{x}_{N+1}) \cdot D_yF(x_N, y_N, z_N) + \beta^N \cdot \mathbb{E}\Big[x_{N+1}D_xF(y_N, y_{N+1}, z_{N+1})|F_{N+1}\Big]\\ &= \beta^N \cdot \tilde{x}_{N+1} \cdot \mathbb{E}\Big[D_xF(y_N, y_{N+1}, z_{N+1})|F_{N+1}\Big]\end{aligned}$$

Now take the expectation of all terms. By smoothing we get

$$\begin{aligned}&\beta^{N-1} \cdot \mathbb{E}\Big[(x_{N+1} - \tilde{x}_{N+1}) \cdot D_yF(x_N, y_N, z_N)\Big] + \beta^N \cdot \mathbb{E}\Big[x_{N+1} \cdot D_xF(y_N, y_{N+1}, z_{N+1})\Big]\\ &= \beta^N \cdot \mathbb{E}\Big[\tilde{x}_{N+1} \cdot \mathbb{E}\Big[D_xF(y_N, y_{N+1}, z_{N+1})|F_{N+1})\Big]\Big].\end{aligned}$$

Now take the limits as $N \to \infty$ (note: $y_N = x_{N+1}$)

$$\begin{aligned}&\lim_{N\to\infty}\Big[\beta^{N-1} \cdot \mathbb{E}\Big[(x_{N+1} - \tilde{x}_{N+1}) \cdot D_yF(x_N, y_N, z_N)\Big]\Big] + \lim_{N\to\infty}\Big[\beta^N \cdot \mathbb{E}\Big[x_{N+1} \cdot D_xF(x_{N+1}, y_{N+1}, z_{N+1})\Big]\Big]\\ &= \lim_{N\to\infty}\Big[\beta^N \cdot \mathbb{E}\Big[\tilde{x}_{N+1} \cdot \mathbb{E}\Big[D_xF(x_{N+1}, y_{N+1}, z_{N+1})F_{N+1}\Big]\Big]\Big]\end{aligned}$$

Since Π satisfies transversality condition, we have

$$\lim_{N\to\infty}\Big[\beta^N \cdot \mathbb{E}\Big[x_{N+1} \cdot D_xF(x_{N+1}, y_{N+1}, z_{N+1})\Big]\Big] = 0$$

Then we have

$$\begin{aligned}&\lim_{N\to\infty}\left[\beta^{N-1}\cdot\mathbb{E}\left[(x_{N+1}-\tilde{x}_{N+1})\cdot D_yF(x_N,y_N,z_N)\right]\right]\\&=\lim_{N\to\infty}\left[\beta^{N}\cdot\mathbb{E}\left[\tilde{x}_{N+1}\cdot\mathbb{E}\left[D_xF(x_{N+1},y_{N+1},z_{N+1})|F_{N+1}\right]\right]\right]\end{aligned}$$

From our assumptions we have that $F(\cdot,\cdot,\cdot)$ is non-decreasing in its first argument, $\tilde{x}_{N+1}\in\mathbb{R}^{\ell}_{+}$. Then $D_xF(x_{N+1},y_{N+1},z_{N+1})\geq 0$ and $\tilde{x}_{N+1}\geq 0$ i.e. all the elements of vector $\tilde{x}_{N+1}$ are equal or greater then zero. By these assumptions we conclude that the right-hand side of the previous equality is equal to or greater than zero. Then

$$\lim_{N\to\infty}\left[\beta^{N-1}\cdot\mathbb{E}\left[(x_{N+1}-\tilde{x}_{N+1})\cdot D_yF(x_N,y_N,z_N)\right]\right]\geq 0$$

In previous parts we showed the following,

$$E(Q_N)-E(\tilde{Q}_N)\geq\beta^{N-1}\cdot\mathbb{E}\left[(x_{N+1}-\tilde{x}_{N+1})\cdot D_yF(x_N,y_N,z_N)\right]$$

Now take the limit as $N\to\infty$.

$$\lim_{N\to\infty}\left[E(Q_N)-E(\tilde{Q}_N)\right]\geq\lim_{N\to\infty}\left[\beta^{N-1}\cdot\mathbb{E}\left[(x_{N+1}-\tilde{x}_{N+1})\cdot D_yF(x_N,y_N,z_N)\right]\right]\geq 0$$

But in previous part we showed that the right-hand side of this inequality is equal to or greater than zero. So we have

$$\lim_{N\to\infty}\left[E(Q_N)-E(\tilde{Q}_N)\right]\geq 0$$

Then we have

$$\lim_{N\to\infty}E(Q_N)\geq\lim_{N\to\infty}E(\tilde{Q}_N)$$

$$\boxed{R^{\Pi}(x_1,z_1)\geq R^{\tilde{\Pi}}(x_1,z_1)}$$

This means that the plan Π which satisfies both Euler equation and transversality condition is better than other plan $\tilde{\Pi}$. Then Π is the optimal plan. (It is assumed that $X\subseteq\mathbb{R}^{\ell}_{+}$). Proof is done.

For a dynamic programming problem (S,A,q,r,β) where S is a countable, non-empty set of possible states, $A(s)$ is the set of actions for $s\in S$, q is the law of motion which decides the probability distribution of the next state based on s,a i.e. $q(\cdot|s,a)$, $r(\cdot,\cdot)$ is a non-negative daily reward function defined on the domain (s,a), $\beta\in(0,1)$ is the discount factor.

With the above technical assumptions and by Theorem 1p, we have that a plan Π is optimal at $s\in S$ if and only if Π is both thrifty and equalizing at $s\in S$. This statement holds certainly only if we restrict our question with these technical assumptions. Now take a look at these technical assumptions. First look at the assumption that S is countable. From probability theory we have that probability space is defined on Borel sets. When taking mean values we use this probability space defined on a Borel set. When S is uncountable, some functions of interest may become "not Borel measurable". When these functions are not Borel measurable, we can not perform our mean value calculations because probability space is defined on a Borel set. When these functions are not Borel measurable, it is impossible to measure a function which is outside the measurable space of the probability space. So the assumption that S is countable is very reasonable and not restrictive. In almost all areas of dynamic programming this assumption is made to reach useful results.

Secondly take a look at the technical assumption that $r(\cdot,\cdot)$ is non-negative. This assumption is made in order to make the sequence $\{\beta^{n-1}\cdot V(s_n)\cdot F_n\}_{n\geq 1}$ a supermartingale hence to make this sequence non-increasing in expectation. Also since it is assumed $r(\cdot,\cdot)$ is non-negative, then this sequence can not be smaller than zero. So by this assumption we have that $\{\beta^{n-1}\cdot V(s_n)\cdot F_n\}_{n\geq 1}$ is non-increasing in expectation and greater than zero in expectation. So this sequence must converge. So the assumption $r(\cdot,\cdot)\geq 0$ is made in order to make this sequence convergent hence to deduce Theorem 1p from the chain of equality and inequalities (2.6) in [0]. But the assumption that $r(\cdot,\cdot)\geq 0$ is a serious drawback because it severely limits the amount of examples we can solve by this model. The maximization of the total discounted reward where the daily reward is non-negative has a very limited application. But making this assumption was crucial in the paper to go through the results. The assumption that $\beta \in (0,1)$ is a general assumption of discounted models hence does not limit the generality of the result.

Now look at the assumptions of the section 3. In this part, it is assumed that $S = X \times Z$ and X is convex. The daily reward function $F(x,y,z)$ is concave in the pair (x,y). Π is interior plan. $F(x,y,z)$ is continuously differentiable in the interior of $(X \times X)$. $\Gamma(x,z)$ and $F(x,y,z)$ are non-decreasing in x. It is assumed that for every interior state (x_0,z) and interior action $y \in \Gamma(x_0,z)$ there exists an open neighbourhood $O \subset X$ of x_0 such that $y \in \Gamma(x,z)\ \forall x \in O$. $\Gamma(x,z)$ is convex in x. So the problem of optimization is a concave programming problem.

First look at the assumptions that X is convex set and $F(x,y,z)$ is concave in (x,y). This assumption is made in order to exploit the properties of concave programming. Concave programming has very useful deductions that were used in this paper to go through the results. But this assumptions are very restrictive and does not hold in general. It means that if our set X is not convex and $F(x,y,z)$ is not concave in (x,y) then we can not use this theorem, which is the case for most of the real-life problems. But in this set-up of the paper the exploitation of concave programming properties was a must. The convexity of $\Gamma(x,z)$ is necessary because current action becomes next state in our model. So $F(x_{n+1},y_{n+1},z_{n+1}) = F(y_n,y_{n+1},z_{n+1})$ where $x_{n+1} \in \Gamma(x_n,z_n)$. So in order for our problem to continue to act like a concave programming problem $x_{n+1} \in \Gamma(x_n,z_n)$ must constitute a convex set So it is assumed that $\Gamma(x,z)$ is convex in x for the continuation of concave programming property over time n.

Secondly look at the assumptions that Π is interior, $F(x,y,z)$ is continuously differentiable in the interior of (X,X). Since $F(x,y,z)$ is defined to be continuously differentiable on the interior of (X,X), Π is interior to hold the domain of the $F(\cdot,\cdot,\cdot)$ in the interior. If Π is not interior then the states can leave the interior, so we can not use continuous differentiability. So Π is interior for the continuation of the domain of $F(\cdot,\cdot,\cdot)$ to be interior. Interiority and continuity assumptions are used in equation (3.5) on [0].

Property 6r: If a function $f(\cdot)$ is continuously differentiable then $f(\cdot)$ is Lipschitz continuous. So if $f(\cdot)$ is continuously differentiable then $f(\cdot)$ has bounded first derivative because Lipschitz continuous functions are limited by some number that how fast they can change.

$$\text{Continuously differentiable} \subset \text{Lipschitz continuous} \quad [6]$$

We used Property 6r when showing Lemma 7 (ii) on [0]. We used concavity for showing monotonicity of derivative and used continuous differentiability to show bounded derivative. Then used monotone convergence because we had both boundedness and monotonicity. So continuous differentiability property is a must since every other proofs onward depends on the result of Lemma 7 (ii) on [0]. Continuous differentiability of $F(x,y,z)$ is needed for bounded first derivative of $V(\cdot,z_n)$ to prove this Lemma. Since $F(x,y,z)$ is assumed to be continuously differentiable by Property 6r it is Lipschitz continuous which means bounded first derivative. If this was not the case then we could not have proven Lemma 7 (ii) on [0] and not be able to derive the other results.

Interiority and continuity assumptions are needed for proving differentiability of $V(\cdot,z_n)$ at z_n. We did this proof by first showing the differentiability of $W(\cdot)$. $W(x)$ is defined as $\Psi(x,y_n,z_n)$.

We know that $W(x) = \Psi(x,y_n,z_n)$ and $W(x_n) = \Psi(x_n,y_n,z_n) = V(x_n,z_n)$. So in order for $W(x)$ to satisfy the properties of $\Psi(\cdot,\cdot,\cdot)$ such as concavity and continuous differentiability, y_n must be in the action space of x i.e. $y_n \in \Gamma(x,z_n)$. So here comes the use of interiority and continuity assumptions. Interiority and continuity implies that there exists, for $\forall n \in N$ on open neighborhood O_n of x_n such that $y_n \in \Gamma(x,z_n)$ holds for $\forall x \in O_n$. Therefore $W(x)$ is concave and continuously differentiable. Since $W(x) = \Psi(x,y_n,z_n)$ if $y_n \in \Gamma(x,z_n)$ the action y_n conserves the properties of $\Psi(x,y_n,z_n)$ in the interior set. These key properties are concavity and continuous differentiability. In order y_n to be in the action set $\Gamma(x,z_n)$, we need continuity and interiority. Then by using concavity and continuous differentiability of $W(\cdot)$ at x_n, we can show the differentiability of $V(\cdot,z_n)$ at x_n and prove Theorem 8 (a) on [0].

Lastly look at the assumptions that $\Gamma(x,z)$ and $F(x,y,z)$ are non-decreasing in x. Those technical assumption provides us that $V(x,z)$ is non-decreasing in x. By using this and concavity of the $V(\cdot,\cdot)$, we showed Lemma 7 (ii) on [0]. And we went through with this result. All the proofs onwards depends on Lemma 7 (ii) and Lemma 7 (ii) depends on Lemma 7 (i) on [0] and this depends on non-decreasing assumptions of $\Gamma(x,z)$ and $F(x,y,z)$ in x.

Interiority, continuity and continuous differentiability assumptions are reasonable in the sense that they are necessary for the proofs in Section 3. But these assumptions are very restrictive and we can only deal with a very special class of problems with these assumptions. In most of dynamic programming examples the optimal plan can be at the boundaries of the defining set. So we can not use this model in the paper for problems which has optimal plans at the boundaries. Also since we can not check this condition without first finding the optimal plan, this model does not provide a very certain approach for finding the optimal plan. This model is mostly helpful for testing potential plans but not finding a plan from the start. Continuity assumption is also very restrictive. Most of the dynamic programmings are on discrete time but in this paper we only deal with continuous time problems. Also the continuity assumption on the actions y is very restrictive. In some dynamic problems the action y_n may be defined on at only a specific point but not defined in the neighborhood. Even if the state is continuous, a specific action is not necessarily continuously transitive between states, for general problems. The assumption that $F(x,y,z)$ is continuously differentiable in the interior of (X,X) is very restrictive. There is no general rule about the cost (reward) functions of dynamic programming problems. In my opinion, amongst the above cases, the most restrictive one is the interiority of the optimal plan. Since that we do not know this without finding it. This makes this model suitable for testing potential plans. The assumptions that $\Gamma(x,z)$ and $F(x,y,z)$ are non-decreasing in x are also very restrictive and does not hold for a bunch of problems.

To sum up, all these technical assumptions are made in order to prove the results. These assumptions are not made in order to solve a general class of dynamic programming problems. These assumptions are made only for proving the results and for reaching a conclusion. After doing these we can search for problems which satisfy those condition and try to solve them. Especially for economists, most of the economic models satisfies the assumptions of Section 3. So this result are useful for some class of economic problems.

If we assume all these restrictions for both "Euler and transversality" and "thrifty and equalizing" models, then these two models become equivalent for the special class of problems with these restrictions which is the case for some economic models.

Pros and Cons of the "Euler and Transversality" Model:

- Easy to apply in many examples: Optimality can be checked without having to calculate the optimal reward function V.
- Requires assumptions of concavity and smoothness of the daily reward r, convexity of the action set A and interiority of the plan Π.

Question: How much can the special assumptions be relaxed?

d) Section 4: Proofs and Results:

Now the paper tries to answer the above question for the special case of one-dimensional i.e. $\mathbb{R}^\ell$, $\ell = 1$. In one dimensional case, we can replace the Envelope equation and Euler equation with appropriate inequalities and remove the restriction of interiority of plan Π. The aim of this section is by turning equalities to inequalities, we try to remove the interiority assumption on plan Π.

Theorem 8p: (Envelope Inequalities) If the plan Π is thrifty at $s = (x_1, z_1)$ then with probability one under $P^{\Pi,s}$, we have for all $n = 1, 2, \ldots$ the properties

$$D_x^+ V(x_n, z_n) \geq D_x^+ F(x_n, y_n, z_n)$$

$$D_x^- V(x_n, z_n) \leq D_x^- F(x_n, y_n, z_n) \text{ on } \{y_n < \delta(x_n, z_n)\}$$

Proof: By thriftiness of plan Π, the actions y_n conserve $V(\cdot,\cdot)$ at (x_n, z_n). So we have $W(x_n) = V(x_n, z_n)$. Since $y_n \in \Gamma(x_n, z_n) \subseteq \Gamma(x_n + \varepsilon, z_n)$ holds for $\varepsilon > 0$ by non-decreasing property we have $W(x_n + \varepsilon) \leq V(x_n + \varepsilon, z_n)$.

$$D_x^+ W(x_n) = \lim_{\varepsilon \to 0} \frac{W(x_n + \varepsilon) - W(x_n)}{\varepsilon}$$

$$D_x^+ V(x_n, z_n) = \lim_{\varepsilon \to 0} \frac{V(x_n + \varepsilon, z_n) - V(x_n, z_n)}{\varepsilon}$$

Since $V(x_n + \varepsilon, z_n) - W(x_n + \varepsilon) \geq 0$ and $W(x_n) = V(x_n, z_n)$ we get

$$V(x_n + \varepsilon, z_n) - W(x_n + \varepsilon) = V(x_n + \varepsilon, z_n) - V(x_n, z_n) - W(x_n + \varepsilon) + W(x_n)$$

Then we get

$$\boxed{\Big(V(x_n + \varepsilon, z_n) - V(x_n, z_n)\Big) - \Big(W(x_n + \varepsilon) - W(x_n)\Big) \geq 0}$$

Taking limit as $\varepsilon \to 0$ gives,

$$\lim_{\varepsilon \to 0} \left[\frac{(V(x_n + \varepsilon, z_n) - V(x_n, z_n)) - (W(x_n + \varepsilon) - W(x_n))}{\varepsilon} \right] \geq 0$$

Then

$$\lim_{\varepsilon \to 0} \frac{(V(x_n + \varepsilon, z_n) - V(x_n, z_n)}{\varepsilon} - \lim_{\varepsilon \to 0} \frac{W(x_n + \varepsilon) - W(x_n)}{\varepsilon} \geq 0$$

Recalling the definition of $W(\cdot)$,

$$W(x_n + \varepsilon) = F(x_n + \varepsilon, y_n, z_n) + \beta \int_Z V(y_n, t) \cdot q(dt|z_n)$$

$$W(x_n) = F(x_n, y_n, z_n) + \beta \int_Z V(y_n, t) \cdot q(dt|z_n)$$

$$W(x_n + \varepsilon) - W(x_n) = F(x_n + \varepsilon, y_n, z_n) - F(x_n, y_n, z_n)$$

Then we have

$$\lim_{\varepsilon \to 0} \frac{(V(x_n + \varepsilon, z_n) - V(x_n, z_n)}{\varepsilon} - \lim_{\varepsilon \to 0} \frac{F(x_n + \varepsilon, y_n, z_n) - F(x_n, y_n, z_n)}{\varepsilon} \geq 0$$

From this we get

$$D_x^+ V(x_n, z_n) - D_x^+ F(x_n, y_n, z_n) \geq 0 \quad \text{Then}$$

$$\boxed{D_x^+ V(x_n, z_n) \geq D_x^+ F(x_n, y_n, z_n)}$$

Now prove the second statement. From assumptions we have $X = [0, \infty)$ and $\Gamma[0, \delta(x, z))$. It follows from the continuity of the function $\delta(\cdot, z_n)$ that $y_n \in \Gamma(x_n - \varepsilon, z_n)$ for $\varepsilon > 0$ sufficiently small. So the optimal control action at x_n, which is y_n, can be used as control action in some neighborhood of x_n. There exists such a neighborhood for some $\varepsilon > 0$. This is an implication of continuity. Hence for such ε, $W(x_n - \varepsilon) \leq V(x_n - \varepsilon, z_n)$. Then

$$V(x_n - \varepsilon, z_n) - W(x_n - \varepsilon) \geq 0$$

Since $V(x_n, z_n) = W(x_n)$ we get

$$V(x_n - \varepsilon, z_n) - V(x_n, z_n) - W(x_n - \varepsilon) + W(x_n) \geq 0$$

Then

$$V(x_n, z_n) - V(x_n - \varepsilon, z_n) - W(x_n) + W(x_n - \varepsilon) \leq 0$$

Then

$$\Big(V(x_n, z_n) - V(x_n - \varepsilon, z_n)\Big) - \Big(W(x_n) - W(x_n - \varepsilon)\Big) \leq 0$$

By taking limits as $\varepsilon \to 0$ we get

$$\lim_{\varepsilon \to 0} \left[\frac{(V(x_n, z_n) - V(x_n - \varepsilon, z_n)) - (W(x_n) - W(x_n - \varepsilon))}{\varepsilon} \right] \leq 0$$

Then we get

$$\lim_{\varepsilon \to 0} \frac{V(x_n, z_n) - V(x_n - \varepsilon, z_n)}{\varepsilon} - \lim_{\varepsilon \to 0} \frac{W(x_n) - W(x_n - \varepsilon)}{\varepsilon} \leq 0$$

$$W(x_n) - W(x_n - \varepsilon) = F(x_n, y_n, z_n) - F(x_n - \varepsilon, y_n, z_n) \quad \text{Then}$$

$$\lim_{\varepsilon\to 0}\frac{V(x_n,z_n)-V(x_n-\varepsilon,z_n)}{\varepsilon}-\lim_{\varepsilon\to 0}\frac{F(x_n,y_n,z_n)-F(x_n-\varepsilon,y_n,z_n)}{\varepsilon}\leq 0$$

Then we have

$$D_x^- V(x_n,z_n)-D_x^- F(x_n,y_n,z_n)\leq 0 \quad \text{then}$$

$$\boxed{D_x^- V(x_n,z_n)\leq D_x^- F(x_n,y_n,z_n)} \quad \text{on} \quad \{y_n<\delta(x_n,z_n)\}$$

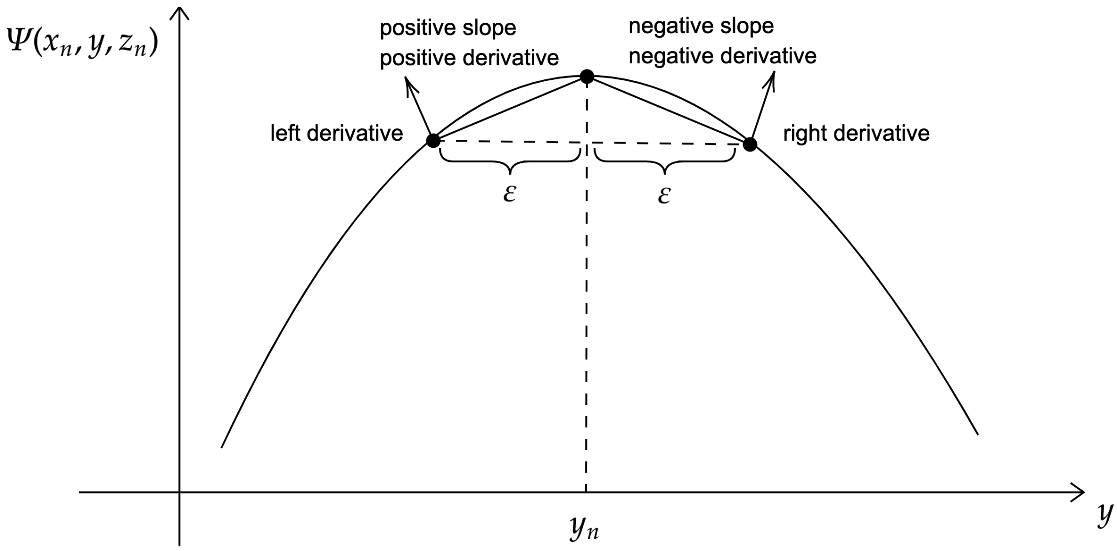
We have the restriction $\{y_n<\delta(x_n,z_n)\}$ because if $y_n\geq\delta(x_n,z_n)$ then it is impossible for y_n be an element of $\Gamma(x_n-\varepsilon,z_n)$. We have this restriction because we need y_n to be an element of action space for some $\varepsilon>0$ at $(x_n-\varepsilon,z_n)$ in order to define the function $W(\cdot)$ as desired.

Theorem 9p: (Euler Inequalities) If the plan Π is thrifty at $s=(x_1,z_1)$ then with probability are under $P^{\Pi,s}$ we have for all $n=1,2,\ldots$ the inequalities

$$0\geq D_y^+F(x_n,y_n,z_n)+\beta\int_Z D_x^+F(y_n,y_{n+1},t)\cdot q(dt|z_n) \quad \text{on} \quad \{y_n<\delta(x_n,z_n)\}$$

$$0\leq D_y^-F(x_n,y_n,z_n)+\beta\int_Z D_x^-F(y_n,y_{n+1},t)\cdot q(dt|z_n) \quad \text{on} \quad \{y_n>0\}$$

Proof: By the thriftiness of Π, the action y_n conserves $V(\cdot,\cdot)$ at each state $s_n=(x_n,z_n)$ $n=1,2,\ldots$. The concave function $\Psi(x_n,y,z_n)$ is maximized over $\Gamma(x_n,z_n)$ at y_n. Let us show what happens with a simple figure



As seen in the figure, the slope of the line which connects y_n and a right side point is negative. If this slope is negative for an arbitrary ε, then the slope will be negative as $\varepsilon\to 0$. So the right hand derivative is negative. The same argument goes for left hand derivative and slope being negative. This implies

$$0\geq D_y^+\Psi(x_n,y_n,z_n) \quad \text{provided} \quad y_n<\delta(x_n,z_n)$$

Here the restriction $y_n<\delta(x_n,z_n)$ is needed because when we are taking right derivative we are getting closer to y_n by right side i.e. as $y_n+\varepsilon\to y_n$ for $\varepsilon>0$. So if $y_n\geq\delta(x_n,z_n)$ then it is impossible for $y_n+\varepsilon$ to be in the action set $\Gamma(x_n,z_n)$ because $y_n+\varepsilon\notin[0,\delta(x_n,z_n))$. Also the figure about concavity implies,

$$0\leq D_y^-\Psi(x_n,y_n,z_n) \quad \text{provided} \quad y_n>0.$$

Here the restriction $y_n>0$ is needed because when we are taking left derivative we are getting closer to y_n by the left side i.e. as $y_n-\varepsilon\to y_n$ for $\varepsilon>0$. So if $y_n\leq 0$ then it is impossible for $y_n-\varepsilon$ to be in the action set $\Gamma(x_n,z_n)$ because $y_n-\varepsilon\notin[0,\delta(x_n,z_n))$. For each $(y,z)\in[0,\infty)XZ$, the quotients $(V(y+\varepsilon,z)-V(y,z))/\varepsilon$, $\varepsilon\geq 0$ are non-negative by Lemma 3p (i). These quotients increase as $\varepsilon\downarrow 0$, as shown in Lemma 3p (ii) by the concavity and non-decreasing property of $V(\cdot, z)$. Also from continuous differentiability of $F(x, y, z)$ we have that $V(\cdot, z)$ is continuously differentiable. This implies that $V(\cdot, z)$ is Lipschitz continuous. For a continuous and differentiable function $V(\cdot, z)$, Lipschitz continuity implies bounded first derivative for $V(\cdot, z)$. With the knowledge of non-decreasing quotient and bounded first derivative we conclude that the quotients as $\varepsilon \downarrow 0$ is convergent because the sequence is both monotone and bounded. By monotone convergence theorem we can show the following

$$D_x^+ \int\limits_Z V(y,t) \cdot q(dt|z) = \int\limits_Z D_x^+ V(y,t) \cdot q(dt|z)$$

also

$$D_x^- \int\limits_Z V(y,t) \cdot q(dt|z) = \int\limits_Z D_x^- V(y,t) \cdot q(dt|z)$$

$$\Psi(x_n, y_n, z_n) = F(x_n, y_n, z_n) + \beta \int\limits_Z V(y_n, t) \cdot q(dt|z_n)$$

$$D_y^+ \Psi(x_n, y_n, z_n) = D_y^+ F(x_n, y_n, z_n) + \beta D_x^+ \int\limits_Z V(y_n, t) \cdot q(dt|z)$$

By monotone convergence of quotients we get,

$$D_y^+ \Psi(x_n, y_n, z_n) = D_y^+ F(x_n, y_n, z_n) + \beta \int\limits_Z D_x^+ V(y_n, t) \cdot q(dt|z)$$

Since $0 \geq D_y^+ \Psi(x_n, y_n, z_n)$ we get

$$0 \geq D_y^+ F(x_n, y_n, z_n) + \beta \int\limits_Z D_x^+ V(y_n, t) \cdot q(dt|z_n)$$

By Envelope equality we have

$$D_x^+ V(y_n, t) \geq D_x^+ F(y_n, y_{n+1}, t)$$

By using this we get

$$D_y^+ F(x_n, y_n, z_n) + \beta \int\limits_Z D_x^+ V(y_n, t) \cdot q(dt|z_n) \geq D_y^+ F(x_n, z_n, y_n) + \beta \int\limits_Z D_x^+ F(y_n, y_{n+1}, t) \cdot q(dt|z_n)$$

So we get

$$\boxed{0 \geq D_y^+ F(x_n, y_n, z_n) + \beta \int\limits_Z D_x^+ F(y_n, y_{n+1}, t) \cdot q(dt|z_n)}$$

The proof for left derivative case is similar. It is now possible in the special setting of this section that a transversality condition is necessary for a plan to be optimal, even without interiority or smoothness of the daily reward.

Theorem 10p: (Transversality Condition) If Π is optimal at $s = (x_1, z_1)$ then

$$\lim_{n \to \infty} \left(\beta^n \cdot \mathbb{E}^{\Pi, s}[x_n \cdot D_x^+ F(x_n, y_n, z_n)] \right) = 0$$

Proof: By ([7] page 13) We have the following

- An elementary argument involving only the concepts of the Rieman integral can be used to show that if f is concave on (a, b) then for all $c, x \in (a, b)$

$$f(x) - f(c) = \int\limits_c^x f'_-(t)\, dt = \int\limits_c^x f'_+(t)\, dt \quad \text{Then,}$$

$$V(x_n, z_n) - V(0, z_n) = \int_0^{x_n} D_x^+ V(x, z_n)\, dx$$

Now start the proof

$$V(x_n, z_n) \geq V(x_n, z_n) - V(0, z_n)$$

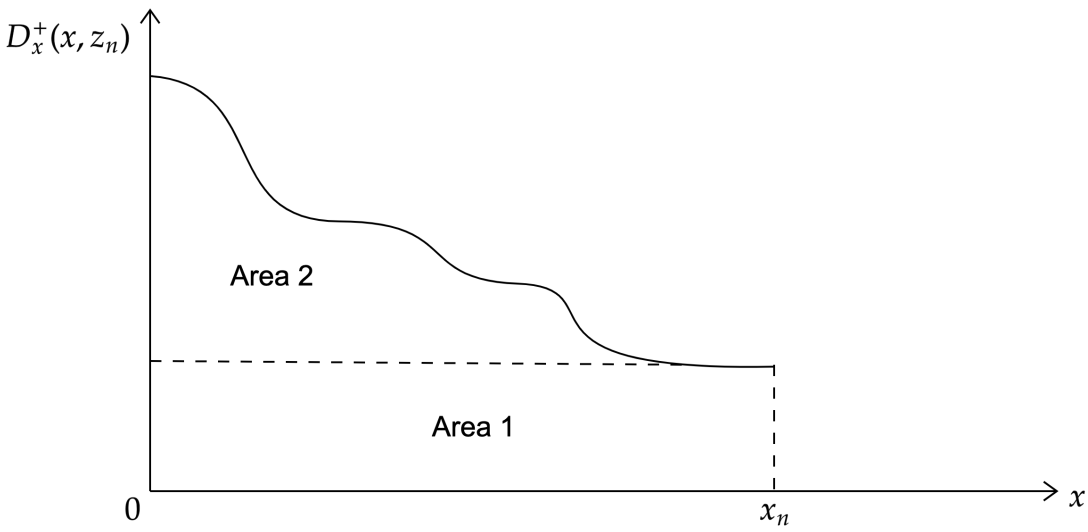
This inequality holds since that $V(0, z_n) \geq 0$. And this is because of the non-negativity of the daily reward.

$$V(x_n, z_n) - V(0, z_n) = \int_0^{x_n} D_x^+ V(x, z_n)\, dx$$

This equality holds by the previous theorem on ([7] page 13). Now try to show

$$\int_c^{x_n} D_X^+ V(x, z_n)\, dx \geq x_n \cdot D_x^+ V(x_n, z_n)$$

First note that since $V(x, z)$ is concave in x, then its right hand side derivative i.e. $D_x^+ V(x, z)$ is decreasing in x. Also we have from assumptions that $V(\cdot, \cdot)$ is non-negative and non-decreasing. So $D_x^+ V(x, z) \geq 0$. Then we have that in the interval $[0, x_n]$, $D_x^+ V(x, z_n)$ is maximum at $x = 0$ and minimum at $x = x_n$ because $D_x^+ V(x, z_n)$ is non-increasing with x. Now consider the following figure



So $\int_0^{x_n} D_x^+ V(x, z_n)\, dx =$ Area 1 + Area 2.
$x_n \cdot D_x^+ V(x, z_n)\, dx =$ Area 1. So

$$\boxed{\int_0^{x_n} D_x^+ V(x, z_n)\, dx \geq x_n \cdot D_x^+ V(x_n, z_n)}$$

This was not a proof but an illustration on how to use the properties and assumptions of $V(\cdot, \cdot)$ for coming up with a figure which illustrates the behavior of the right derivative of the $V(\cdot, \cdot)$. We do not know the general graph but it will resemble something similar to I draw. Then our inequality follows.

$$x_n \cdot D_x^+ V(x_n, z_n) \geq x_n \cdot D_x^+ F(x_n, y_n, z_n)$$

We have this inequality by Envelope inequality.

$$x_n \cdot D_x^+ F(x_n, y_n, z_n) \geq 0$$

Since $X = [0, \infty)$ and $F(\cdot, y, z)$ is non-decreasing we have this inequality. By this chain of equality and inequalities we get,

$$V(x_n, z_n) \geq x_n \cdot D_x^+ F(x_n, y_n, z_n) \geq 0$$

Now multiply this with β^n. Then we get

$$\beta^n \cdot V(x_n, z_n) \geq \beta^n \cdot x_n \cdot D_x^+ F(x_n, y_n, z_n) \geq 0$$

Now take the limit as $n \to \infty$. Then

$$\lim_{n\to\infty} \left[\beta^n \cdot \mathbb{E}^{\Pi,s}(V(x_n, z_n))\right] \geq \lim_{n\to\infty} \left[\beta^n \cdot \mathbb{E}^{\Pi,s}(x_n \cdot D_x^+ F(x_n, y_n, z_n))\right] \geq 0$$

Since Π is optimal, then Π is equalizing. So we have by equalizing property the following.

$$\lim_{n\to\infty} \left[\beta^n \cdot \mathbb{E}^{\Pi,s}(V(x_n, z_n))\right] = 0 \quad \text{Then we have}$$

$$0 \geq \lim_{n\to\infty} \left[\beta^n \cdot \mathbb{E}^{\Pi,s}(x_n \cdot D_x^+ F(x_n, y_n, z_n))\right] \geq 0$$

This implies

$$\boxed{\lim_{n\to\infty} \left[\beta^n \cdot \mathbb{E}^{\Pi,s}(x_n \cdot D_x^+ F(x_n, y_n, z_n))\right] = 0}$$

An Example: $S = \{(x, y) : \; x \geq 0, \; y \geq 0\}$ where x is cash and y is goods for sale at price p.

$$A(x, y) = \left\{a : \; 0 \leq a \leq x + \frac{py}{1+p}\right\}$$

where "a" is cash spent for consumption and p is interest rate charged or paid by a bank.

$$r((x, y), a) = u(a/p)$$

where $u : \mathbb{R}^+ \to \mathbb{R}^+$ is concave, increasing and differentiable. The law of motion:

$$(x, y) \to ((1 + p) \cdot (x - a) + py, Y) = (\tilde{x}, Y)$$

where Y is random with given distribution.

The Bellman Equation:

$$\boxed{V(x, y) = \sup_{0 \leq a \leq x + \frac{py}{1+p}} \left[\Psi(a)\right]}$$

where $\Psi(a) = u\left(\frac{a}{p}\right) + \beta\mathbb{E}\left(V((1 + p) \cdot (x - a) + py, Y)\right) = u\left(\frac{a}{p}\right) + \beta\mathbb{E}\left(V(\tilde{x}, Y)\right)$

The value function is concave, increasing and differentiable. Suppose the max is achieved at an interior action a. Then we have

The Envelope Equation:

$$\frac{\partial V}{\partial x}(x, y) = \frac{1}{p} u'\left(\frac{a}{p}\right)$$

Euler and Thriftiness:

With $\Psi = u\left(\frac{a}{p}\right) + \beta\mathbb{E}\left(V((1 + p) \cdot (x - a) + py, Y)\right) = u\left(\frac{a}{p}\right) + \beta\mathbb{E}\left(V(\tilde{x}, Y)\right)$ we can set $\Psi' = 0$ and use the envelop equation to get

The Euler Equation:

$$\frac{1}{p} u'\left(\frac{a}{p}\right) = \frac{\beta(1 + p)}{p} \mathbb{E}\left(u'(\frac{\tilde{a}}{p})\right)$$

where $\tilde{a}$ is the optimal action on the next day.

Theorem: A plan Π that uses interior actions is thrifty if and only if the Euler equation holds almost surely.

The Transversality Condition (TC): The plan Π satisfies TC at an initial state $s = (x, y)$ if

$$\lim_{n\to\infty}\left[\beta^{n-1}\cdot\mathbb{E}^{\Pi,s}\left(x_n\cdot u'(\frac{a_n}{p})\right)\right] = 0$$

Theorem: If Π is an optimal interior plan, then it satisfies TC.

$$V(x_n, y_n) \geq V(x_n, y_n) - V(0, y_n) \geq x_n\cdot\frac{\partial V}{\partial x}(x_n, y_n) = x_n\cdot\frac{1}{p}u'\left(\frac{a_n}{p}\right)$$

By the equalizing property,

$$\lim_{n\to\infty}\left[\beta^{n-1}\cdot\mathbb{E}^{\Pi,s}\big(V(x_n, y_n)\big)\right] = 0$$

Theorem: For an interior plan Π, optimality is equivalent to satisfying TC and Euler conditions. See Stokey and Lucas (1989) for a proof. The result holds in greater generality than our example but does assume: concavity and smoothness of the daily reward r, convexity of the action sets A, and that the plan Π is interior.

A Special Case: Suppose that $Y = y$ is constant and $\beta(1 + p) = 1$ (Fisher's equation). Then the Euler equation is

$$u'\left(\frac{a}{p}\right) = u'\left(\frac{\tilde{a}}{p}\right)$$

which holds if $a = \tilde{a}$ and this holds if $x = \tilde{x}$. So we need

$$x = \tilde{x} = (1 + p)\cdot(x - a) + py \quad \text{or} \quad a = \frac{p}{1+p}x + \frac{pp}{1+p}$$

Since $x_n = x$ and $a_n = a$, the TC is

$$\beta^{n-1}\cdot x_n\cdot u'\left(\frac{a_n}{p}\right) = \beta^{n-1}\cdot x\cdot u'\left(\frac{a}{p}\right) \to 0$$

So the plan is optimal.

General dynamic optimization problem

$$\begin{aligned} V(\hat{x}_0) = \max_{\{x_{t+1}\}_{t=0}^{\infty}} &\sum_{t=0}^{\infty}\beta^t U(x_t, x_{t+1}) \quad \text{s.t.} \\ x_{t+1} \in \Gamma(x_t), &\quad x_0 = \hat{x}_0. \end{aligned} \tag{P}$$

The following are necessary and sufficient conditions for $\{x_{t+1}\}_{t=0}^{\infty}$ to be optimal: if x_{t+1} is in the interior of $\Gamma(x_t)$

$$U_y(x_t, x_{t+1}) + \beta U_x(x_{t+1}, x_{t+2}) = 0, \quad \forall t \tag{EE}$$

$$\lim_{T\to\infty}\beta^T U_y(x_T, x_{T+1})\cdot x_{T+1} = 0 \tag{TC}$$

and $x_0 = \hat{x}_0$.

(EE) together with (TC) and initial condition $x_0 = \hat{x}_0$ fully characterizes optimal $\{x_{t+1}\}_{t=0}^{\infty}$.

(EE) is called "Euler equation" simply first-order condition (FOC) w/ respect to x_{t+1} derivation: differentiate problem (P) with respect to x_{t+1}.

(TC) is called "transversality condition" understanding it is harder than (EE). Stokey-Lucas-Prescott write (TC) as

$$\lim_{T\to\infty}\beta^T U_x(x_T, x_{T+1})\cdot x_T = 0 \tag{TC}$$

Example : Growth Model
Social Planner's problem in growth model

$$
\begin{aligned}
V(\hat{k}_0) = \max_{\{k_{t+1}\}_{t=0}^{\infty}} \sum_{t=0}^{\infty} \beta^t u(f(k_t) + (1-\delta)k_t - k_{t+1}) \quad \text{s.t.} \\
k_{t+1} \in [0, f(k_t) + (1-\delta)k_t], \quad k_0 = \hat{k}_0.
\end{aligned}
\tag{P'}
$$

(EE) and (TC) are

$$
-u'(f(k_t) + (1-\delta)k_t - k_{t+1}) + \beta u'(f(k_{t+1}) + (1-\delta)k_{t+1} - k_{t+2})(f'(k_{t+1}) + 1 - \delta) = 0 \tag{EE'}
$$

$$
\lim_{T\to\infty} \beta^T u'(f(k_T) + (1-\delta)k_T - k_{T+1})k_{T+1} = 0 \tag{TC'}
$$

Get (EE') simply by differentiating (P') with respect to k_{t+1}.

(EE') can be written more intuitively as

$$
\underbrace{\frac{u'(c_t)}{\beta u'(c_{t+1})}}_{MRS} = \underbrace{f'(k_{t+1}) + 1 - \delta}_{MRT}
$$

MRS between c_t and c_{t+1} $= MRT$ between c_t and c_{t+1}

$$
\max_{c_A, c_B} u(c_A, c_B) \quad \text{s.t.} \quad c_A = f(\ell_A),\ c_B = f(\ell_B),\ \ell_A + \ell_B \le 1
$$

where A=apples, B=bananas

$$
\Rightarrow \frac{u_{c_A}(c_A, c_B)}{u_{c_B}(c_A, c_B)} = \frac{f'(\ell_A)}{f'(\ell_B)}
$$

growth model: different dates = different goods.

Summarizing all necessary conditions

$$
\begin{aligned}
u'(c_t) &= \beta u'(c_{t+1})(f'(k_{t+1}) + 1 - \delta) \\
k_{t+1} &= f(k_t) + (1-\delta)k_t - c_t
\end{aligned}
\tag{DE}
$$

for all t, with

$$
\begin{aligned}
k_0 &= \hat{k}_0 \\
\lim_{T\to\infty} \beta^T u'(c_T)k_{T+1} &= 0
\end{aligned}
\tag{TC'}
$$

(DE) is system of two difference equations in $(c_t, k_t) \cdots$

$\cdots$ needs two boundary conditions

1) initial condition for capital stock: $k_0 = \hat{k}_0$
2) transversality condition, plays role of boundary condition

Where does (TC) come from?
Consider finite horizon problem:

$$
\begin{aligned}
V(\hat{k}_0, T) = \max_{\{k_{t+1}\}_{t=0}^{T}} \sum_{t=0}^{t} \beta^t u(c_t) \quad \text{s.t.} \\
k_{t+1} = f(k_t) + (1-\delta)k_t - c_t, \quad k_{t+1} \ge 0.
\end{aligned}
$$

Lagrangean

$$
\mathcal{L} = \sum_{t=0}^{t} \beta^t u(c_t) + \sum_{t=0}^{T} \lambda_t (f(k_t) + (1-\delta)k_t - c_t - k_{t+1}) + \sum_{t=0}^{T} \mu_t k_{t+1}
$$

Necessary conditions at $t = T$

$$\begin{aligned} \beta^T u'(c_T) &= \lambda_T \\ \lambda_T &= \mu_T \qquad \Rightarrow \quad \beta^T u(c_T) k_{T+1} = 0 \\ \mu_T k_{T+1} &= 0 \end{aligned}$$

In the finite horizon

$$\beta^T u'(c_T) k_{T+1} = 0 \tag{$*$}$$

$(*)$ is really two conditions in one

1) $\beta^T u'(c_T) > 0$: need $k_{T+1} = 0$
2) $\beta^T u'(c_T) = 0$: k_{T+1} can be > 0

Intuition for case 1: if my marginal utility of consumption at T is positive, I want to eat up all my wealth before I die

(TC) is same condition as $(*)$ in economy with $T \to \infty$

$$\lim_{T\to\infty} \beta^T u'(c_T) k_{T+1} = 0 \tag{TC}$$

Intuition:

- capital should not grow too fast compared to marginal utility
- e.g. with $u(c) = \log c : \beta^T k_{T+1}/c_{T+1} \to 0$
- if I save too much/spend too little, I'm not behaving optimally

(TC) rules out overaccumulation of wealth.

In practice, often don't have to impose (TC) exactly. Instead, just have to make sure trajectories "don't blow up". E.g. consider growth model: since $\beta < 1$, easy to see that

$$\lim_{T\to\infty} \beta^T u'(c_T) k_{T+1} = 0$$

whenever

$$\lim_{T\to\infty} c_T = c^*, \quad \lim_{T\to\infty} k_T = k^*$$

with $0 < c^*, k^* < \infty$ which is satisfied if $\{c_t, k_t\}$ converge to steady state.

Checking If These Conditions are Satisfied:

1) **Euler Equation and Transversality Condition:**
Checking if the Euler equation holds is just a simple mathematical calculation. For a given model, we already know daily reward $F(x, y, z)$ and transition probabilities $q(dt|z)$. Then we need only to take derivatives and integrals and calculate them then check if for a given plan they are equal to zero. If at some time n, the Euler equation is not zero the plan is not optimal.
Checking if the transversality holds is also a simple mathematical calculation. We already know daily reward and discount factor. We take derivative of $F(x, y, z)$ with respect to x and calculate it at x_n. Then multiply it with x_n. Then for a given plan find the expected value of it then take the limit as $n \to \infty$. IT is just about computational work. So Euler and transversality conditions are very useful in practice because they can be easily applied to find optimal plans. It is only about mathematical calculations.
In Euler and transversality case, assumptions are very restrictive but application is easy. Optimality can be checked without having to calculate optimal reward V.

2) **Thrifty and Equalizing Model:**
This model is not useful for testing if a policy is optimal. This model is useful for showing some properties of dynamic programs and making qualitative arguments about dynamic programs. In real life dynamic programming problems this is not helpful for finding plans, because we need to know the optimal reward V before finding the optimal plan. This makes this model unpractical for real-life

problems. Since thriftiness and equalizing models are unapplicable to real life problems, there is not so much papers and discussion about it. This model is invented not for finding plans but for trying to see some aspects of dynamic programs, and reaching some results. If we have already known the optimal reward V, that would mean that we know the optimal plan. This makes this model useless in practice.

Example a1: Solving a parametric dynamic programming problem. In this example we will illustrate how to solve dynamic programming problem by finding a corresponding value function. Consider the following functional equation:

$$V(k) = \max_{c,k'}\{\log c + \beta V(k')\}$$
$$\text{s.t. } c = Ak^\alpha - k'.$$

The budget constraint is written as an equality constraint because we know that preferences represented by the logarithmic utility function exhibit strict monotonicity - goods are always valuable, so they will not be thrown away by an optimizing decision maker. The production technology is represented by a Cobb-Douglass function, and there is full depreciation of the capital stock in every period:

$$\underbrace{F(k,1)}_{Ak^\alpha 1^{1-\alpha}} + \underbrace{(1-\delta)}_{0} k.$$

A more compact expression can be derived by substitutions into the Bellman equation:

$$V(k) = \max_{k'\geq 0}\{\log[Ak^\alpha - k'] + \beta V(k')\}.$$

We will solve the problem by iterating on the value function. We begin with an initial "guess" $V_0(k) = 0$, that is, a function that is zero-valued everywhere.

$$\begin{aligned} V_1(k) &= \max_{k'\geq 0}\{\log[Ak^\alpha - k'] + \beta V_0(k')\} \\ &= \max_{k'\geq 0}\{\log[Ak^\alpha - k'] + \beta \cdot 0\} \\ &\max_{k'\geq 0}\{\log[Ak^\alpha - k']\}. \end{aligned}$$

This is maximized by taking $k' = 0$. Then $V_1(k) = \log A + \alpha logk$.

Going to the next step in the iteration,

$$\begin{aligned} V_2(k) &= \max_{k'\geq 0}\{\log[Ak^\alpha - k'] + \beta V_1(k')\} \\ &= \max_{k'\geq 0}\{\log[Ak^\alpha - k'] + \beta[\log A + \alpha\log k']\}. \end{aligned}$$

The first-order condition now reads

$$\frac{1}{Ak^\alpha - k'} = \frac{\beta\alpha}{k'} \Rightarrow k' = \frac{\alpha\beta A k^\alpha}{1+\alpha\beta}$$

We can interpret the resulting expression for k' as the rule that determines how much it would be optimal to save if we were at period $T-1$ in the finite horizon model. Substitution implies

$$\begin{aligned} V_2(k) &= \log\left[Ak^\alpha - \frac{\alpha\beta Ak^\alpha}{1+\alpha\beta}\right] + \beta\left[\log A + \alpha\log\frac{\alpha\beta Ak^\alpha}{1+\alpha\beta}\right] \\ &= (\alpha+\alpha^2\beta)\log k + \log\left(A - \frac{\alpha\beta A}{1+\alpha\beta}\right) + \beta\log A + \alpha\beta\log\frac{\alpha\beta A}{1+\alpha\beta}. \end{aligned}$$

We could now use $V_2(k)$ again in the algorithm to obtain a $V_3(k)$, and so on. We know by the characterizations that this procedure would make the sequence of value functions converge to some $V^*(k)$. However, there is a more direct approach, using a pattern that appeared already in our iteration.

Let

$$a \equiv \log\left(A - \frac{\alpha\beta A}{1+\alpha\beta}\right) + \beta\log A + \alpha\beta\log\frac{\alpha\beta A}{1+\alpha\beta}$$

and

$$b \equiv (\alpha + \alpha^2\beta).$$

Then $V_2(k) = a + b \log k$. Recall that $V_1(k) = \log A + \alpha \log k$, i.e., in the second step what we did was plug in a function $V_1(k) = a_1 + b_1 \log k$, and out came a function $V_2(k) = a_2 + b_2 \log k$. This clearly suggests that if we continue using our iterative procedure, the outcomes $V_3(k)$, $V_4(k), \ldots, V_n(k)$, will be of the form $V_n(k) = a_n + b_n \log k$ for all n. Therefore, we may already guess that the function to which this sequence is converging has to be of the form:

$$V(k) = a + b \log k.$$

So let us guess that the value function solving the Bellman has this form, and determine the corresponding parameters a, b:

$$V(k) = a + b \log k = \max_{k' \geq 0}\{\log(Ak^\alpha - k') + \beta(a + \beta \log k')\} \quad \forall\, k.$$

Our task is to find the values of a and b such that this equality holds for all possible values of k. If we obtain these values, the functional equation will be solved. The first-order condition reads:

$$\frac{1}{Ak^\alpha - k'} = \frac{\beta b}{k'} \Rightarrow k' = \frac{\beta b}{1 + \beta b} A k^\alpha.$$

We can interpret $\dfrac{\beta b}{1 + \beta b}$ as a savings rate. Therefore, in this setup the optimal policy will be to save a constant fraction out of each period's income.
Define

$$LHS \equiv a + b \log k$$

and

$$RHS \equiv \max_{k' \geq 0}\{\log(Ak^\alpha - k') + \beta(a + b \log k')\}.$$

Plugging the expression for k' into the RHS, we obtain:

$$\begin{aligned}
RHS &= \log\left(Ak^\alpha - \frac{\beta b}{1 + \beta b} A k^\alpha\right) + a\beta + b\beta \log\left(\frac{\beta b}{1 + \beta b} A k^\alpha\right) \\
&= \log\left[\left(1 - \frac{\beta b}{1 + \beta b}\right) A k^\alpha\right] + a\beta + b\beta \log\left(\frac{\beta b}{1 + \beta b} A k^\alpha\right) \\
&= (1 + b\beta)\log A + \log\left(\frac{1}{1 + b\beta}\right) + a\beta + b\beta \log\left(\frac{\beta b}{1 + \beta b}\right) + (\alpha + \alpha\beta b)\log k.
\end{aligned}$$

Setting $LHS = RHS$, we produce

$$\begin{cases}
a = (1 + b\beta)\log A + \log\left(\dfrac{1}{1 + b\beta}\right) + a\beta + b\beta \log\left(\dfrac{\beta b}{1 + \beta b}\right) \\
b = \alpha + \alpha\beta b
\end{cases}$$

which amounts to two equations in two unknowns. The solutions will be

$$b = \frac{\alpha}{1 - \alpha\beta}$$

and, using this finding,

$$a = \frac{1}{1 - \beta}[(1 + b\beta)\log A + b\beta \log(b\beta) - (1 + b\beta)\log(1 + b\beta)],$$

so that

$$a = \frac{1}{1 - \beta}\frac{1}{1 - \alpha\beta}[\log A + (1 - \alpha\beta)\log(1 - \alpha\beta) + \alpha\beta \log(\alpha\beta)].$$

Going back to the savings decision rule, we have:

$$\begin{aligned}
k' &= \frac{b\beta}{1 + b\beta} A k^\alpha \\
k' &= \alpha\beta A k^\alpha.
\end{aligned}$$

If we let y denote income, that is, $y \equiv Ak^{\alpha}$, then $k' = \alpha\beta y$. This means that the optimal solution to the path for consumption and capital is to save a constant fraction $\alpha\beta$ of income. This setting provides a justification to a constant savings rate. It is a very special setup however, one that is quite restrictive in terms of functional forms.

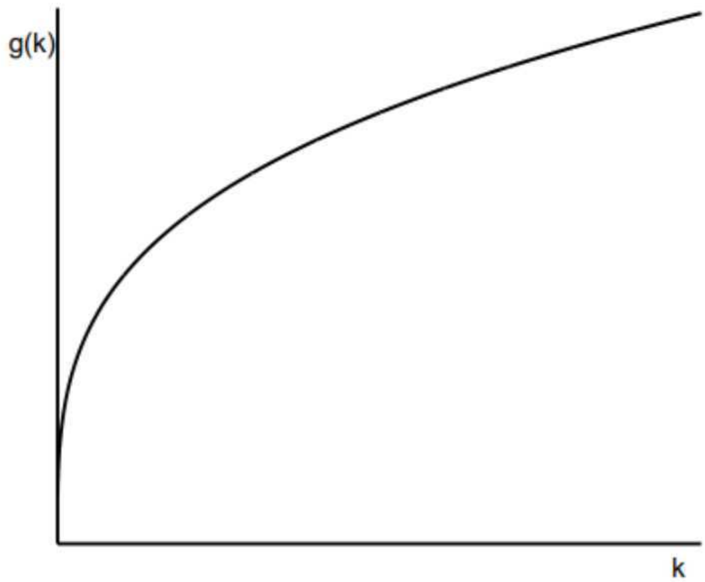


Figure : The decision rule in our parameterized model

The functional Euler equation In the sequentially formulated maximization problem, the Euler equation turned out to be a crucial part of characterizing the solution. With the recursive strategy, an Euler equation can be derived as well. Consider again

$$V(k) = \max_{k' \in \Gamma(k)} \{F(k, k') + \beta V(k')\}.$$

As already pointed out, under suitable assumptions, this problem will result in a function $k' = g(k)$ that we call decision rule, or policy function. By definition, then, we have

$$V(k) = F(k, g(k)) + \beta V[(g(k)].$$

Moreover, $g(k)$ satisfies the first-order condition

$$F_2(k, k') + \beta V'(k') = 0,$$

assuming an interior solution. Evaluating at the optimum, i.e., at $k' = g(k)$, we have

$$F_2(k, g(k)) + \beta V'(g(k)) = 0.$$

One problem in our characterization is that $V'(\cdot)$ is not known: in the recursive strategy, it is part of what we are searching for. However, although it is not possible in general to write $V(\cdot)$ in terms of primitives, one can find its derivative. One can differentiate both sides with respect to k, since the equation holds for all k and, again under some assumptions stated earlier, is differentiable. We obtain

$$V'(k) = F_1[k, g(k)] + \underbrace{g'(k)\{F_2[k, g(k)] + \beta V'[g(k)]\}}_{\text{indirect effect through optimal choice of } k'}$$

From the first-order condition, this reduces to

$$V'(k) = F_1[k, g(k)],$$

which again holds for all values of k. The indirect effects thus disappears: this is an application of the envelope theorem.

Updating, we know that $V'[g(k)] = F_1[g(k), g(g(k))]$ also has to hold. The first order condition can now be rewritten as follows:

$$F_2[k, g(k)] + \beta F_1[g(k),\, g(g(k))] = 0 \quad \forall k.$$

This is the Euler equation stated as a functional equation: it does not contain the unknowns k_t, k_{t+1}, and k_{t+2}. Recall Euler equation formulation

$$F_2[k_t, k_{t+1}] + \beta F_1[k_{t+1}, k_{t+2}] = 0, \quad \forall\, t,$$

where the unknown was the sequence $\{k_t\}_{t=1}^{\infty}$. Now instead, the unknown is the function g. That is, under the recursive formulation, the Euler Equation turned into a functional equation.

The previous discussion suggests that a third way of searching for a solution to the dynamic problem is to consider the functional Euler equation, and solve it for the function g. We can (i) look for sequences solving a nonlinear difference equation plus a transversality condition; or (ii) we can solve a Bellman (functional) equation for a value function.

The functional Euler equation approach is, in some sense, somewhere in between the two previous approaches. It applies more structure than our previous Euler equation. There, a transversality condition needed to be invoked in order to find a solution. Here, we can see that the recursive approach provides some extra structure: it tells us that the optimal sequence of capital stocks needs to be connected using a stationary function.

One problem is that the functional Euler equation does not in general have a unique solution for g. It might, for example, have two solutions. This multiplicity is less severe, however, than the multiplicity in a second-order difference equation without a transversality condition: there, there are infinitely many solutions.

The functional Euler equation approach is often used in practice in solving dynamic problems numerically.

Example a2: In this example we will apply functional Euler equation described above to the model given in Example a1. The function $F(\cdot,\cdot)$ has the form

$$F(k, k') = u(f(k) - g(k))$$

Then, the respective derivatives are:

$$F_1(k, k') = u'(f(k) - k')f'(k)$$
$$F_2(k, k') = -u'(f(k) - k').$$

In the particular parametric example,

$$\frac{1}{Ak^\alpha - g(k)} - \frac{\beta\alpha A(g(k))^{\alpha-1}}{A(g(k))^\alpha - g(g(k))} = 0,\ \forall k.$$

This is a functional equation in $g(k)$. Guess that $g(k) = sAk^\alpha$, i.e. the savings are a constant fraction of output. Substituting this guess into functional Euler equation delivers:

$$\frac{1}{(1-s)Ak^\alpha} = \frac{\alpha\beta A(sAk^\alpha)^{\alpha-1}}{A(sAk^\alpha)^\alpha - sA(sAk^\alpha)^\alpha}.$$

As can be seen, k cancels out, and the remaining equation can be solved for s. Collecting terms and factoring out s, we get

$$s = \alpha\beta.$$

This is exactly the answer that we got in Example a1.

Follow Up Work:

1) Inflationary Equilibrium in a Stochastic Economy with Independent Agents by John Geanakoplos, Ioannis Karatzas, Martin Shubik, William D. Sudderth. [8]
In page 5, Lemma 4.1 part (c) of this paper, a result from this paper [0] is used. In the proof of Lemma 4.1 part (c) the characterization that is constructed in [0] is used. Proof is done by the results of [0]. The statement: "Thriftiness is equivalent to the condition that Π selects actions which attain the supremum in the Bellman equation" from the [0] is used in this Lemma 4.1 part (c).

2) Finitely Additive Dynamic Programming by William D. Sudderth. [9]
In page 104 Remark 7.1 a result from [0] is used. The result used from [0] is the following. "A plan Π is thrifty at s if it chooses actions that achieve the supremum in the Bellman equation and equalizing at s if

$$\lim_{n\to\infty}\left[\beta^n \cdot \mathbb{E}^{\Pi,s}[V(s_{n+1})]\right] = 0$$

This result from [0] is used together with characterization of optimality from [1]. For the proofs in Finitely Additive Dynamic Programming, original characterization on [1] of thrifty and equalizing policies is used together with the further results of [0].

3) Maximizing the Probability of Attaining a Target Prior to Extinction by Debasish Chatterjee, Eugenio Cinquemani and John Lygeros. [10]
In this paper, results about martingale properties and their implications from [0] is used. Also thrifty and equalizing properties and their implications are used in this paper. The general methodology of [0] is also adapted for this paper in order for derivation of optimal policies.

4) Subjective Information Choice Processes by David Dillenberger, R. Vijay Krishna, Philipp Sadowski. [11]
In the proof of Lemma E.14 of this paper characterization of optimal plans from [0] is used. In this paper, in order to test the optimality of σ in Lemma E.14, it is checked that if σ is both conserving and equalizing, which is exactly the result of [0].

References

[0] Karatzas, Ioannis and William Sudderth. 2010. *"Two Characterizations of Optimality in Dynamic Programming"*. Applied Mathematics and Optimization 61: 421-434

[1] Dubins, L. E., Savage, L. J. : *How to Gamble if You Must: Inequalities for Stochastic Processes*. McGraw-Hill, New York (1965).

[2] Stokey, N. L., Lucas, R. E. : *Recursive Methods in Economic Dynamics*. Harvard University Press, Cambridge (1989).

[3] Blackwell, D., *On Stationary Policies*. J. R. Staat. Soc. A 133, 33-37 (1970).

[4] Hordijk, A., *Dynamic Programming and Markov Potential Theory*. Mathematical Centre Tracts 51, Amsterdam (1974).

[5] Rieder, U., *On optimal policies and martingales in dynamic programming*. J. Appl. Probab. 13, 507-518 (1976)

[6] Wikipedia

[7] Robert, A. W. Varberg, D. E., *Convex Functions*. Academic Press, New York (1973)

[8] John Geanakoplos, Ioannis Karatzas, Martin Shubik, and William D. Sudderth, 2009. *"Inflationary Equilibrium In a Stochastic Economy with Independent Agents"*. Cowles Foundation Discussion Papers 1708, Cowles Foundation for Research in Economics, Yale University.

[9] William D. Sudderth, 2016. *"Finitely Additive Dynamic Programming."* Mathematics of Operations Research, INFORMS, vol. 41(1), pages 92-108, February.

[10] D. Chatterjee, E. Cinquemani, and J. Lygeros. Maximizing the probability of attaining a target prior to extinction. *Nonlinear Analysis: Hybrid Systems*, 5(2):367–381, 2011.

[11] D. Dillenberger, R. V. Krishna, and P. Sadowski. *Subjective information choice processes*. University of Pennsylvania and Florida State University and Duke University, 2018.