Lecture 10

I present a specific example of the growth model discussed at the end of lecture 9.

Example 10.1: Let \( y = F(K) = 2 \sqrt{K} \), \( v(C) = \ln(C) \), and \( 0 < \beta < 1 \).

The objective is to maximize \( \sum_{t=0}^{\infty} \beta^t \ln(C_t) \) subject to \( C_t + K_t \leq 2 \sqrt{K_{t+1}} \), for \( t = 0, 1, 2, \ldots \), where \( K_t \) is given. It is easy to derive a first order condition for this problem. For fixed \( t \), hold \( C_t \) constant for \( 0 \leq s < t \) and for \( s > t + 1 \) and hold \( K_s \) constant for \( s \neq t \). The terms of the objective function involving \( C_t \) and \( C_{t+1} \) are \( \beta^t \ln(C_t) + \beta^{t+1} \ln(C_{t+1}) \), which may be written as

\[ g(K_t) = \beta^t \ln(2 \sqrt{K_{t+1}} - K_t) + \beta^{t+1} \ln(2 \sqrt{K_t} - K_{t+1}) \].

The first order condition for a maximum of \( g \) is

\[ 0 = \frac{dg(K_t)}{dK_t} = -\frac{\beta^t}{2 \sqrt{K_{t+1}} - K_t} + \frac{\beta^{t+1}}{2 \sqrt{K_t} - K_{t+1}} \cdot \frac{1}{\sqrt{K_t}} \]

which may be written as

\[ \frac{\beta^{t+1}}{C_{t+1}} \frac{1}{\sqrt{K_t}} = \frac{\beta^t}{C_t} \]

which in turn is the same as

\[ \frac{1}{C_t} = \beta \frac{1}{C_{t+1}} \frac{1}{\sqrt{K_t}} \]

This equation is called the Euler equation for the problem. It may be written as

\[ C_{t+1} = \beta \frac{C_t}{\sqrt{K_t}} \quad (10.1) \]

This equation plus the feasibility equation

\[ C_{t+1} + K_{t+1} = F(K_t) = 2 \sqrt{K_t} \]

determine the evolution of the program over time, for suppose we know \( C_t \) and \( K_t \). Then \( C_{t+1} \) is determined by the Euler equation 10.1 and \( K_{t+1} \) is then determined by the feasibility equation.
In order to find a feasible program satisfying these equations, assume that

\[ K_t = sY_t = sF(K_{t-1}) = s2\sqrt[4]{K_{t-1}}, \quad (10.2) \]

for all \( t \geq 0 \), where \( 0 < s < 1 \), \( s \) is independent of \( t \), and \( Y_t = 2\sqrt[4]{K_{t-1}} \) is output in period \( t \). The parameter \( s \) is the proportion of output that is saved and invested. For \( t = 0, 1, 2, \ldots \),

\[ C_t = (1 - s)2\sqrt[4]{K_{t-1}}, \quad (10.3) \]

since \( C_t + K_t = 2\sqrt[4]{K_{t-1}} \). By substituting equation 10.3 into equation 10.1 and simplifying, we obtain the equation

\[ K_t = \beta\sqrt[4]{K_{t-1}}. \quad (10.4) \]

This equation gives us the value of the constant \( s \) appearing in equation 10.2, namely,

\[ s = \beta/2. \]
The difference equation 10.4 converges, and the limit is $\beta^2$. That is, if $K_t$ evolves according to equation 10.4 then

$$\lim_{t \to \infty} K_t = \beta^2.$$ 

The convergence may be seen in the figure on the previous page. The straight line in the diagram is the diagonal $K_t = K_{t-1}$. The curve, labeled $\beta \sqrt{K_t}$ is the graph of this function.

Because the sequence $K_t$ converges, so does the sequence $C_t$, and its limit is a positive number, namely $2\beta - \beta^2 = \beta(2 - \beta)$.

I argue that the program generated in this way by $s = \beta/2$ is optimal. Notice that the initial capital stock, $K_0$, is determined by the equation

$$K_0 = sy_0 = \frac{\beta y_0}{2},$$

where $y_0$ is the initial stock of output.
I explore what happens if we start with a value of $K_0$ different from $\frac{\beta}{2}$, and allow the program to be generated by the Euler equation and feasibility. Define the number $a_t$ by the equation

$$C_t = a_t 2^{\frac{K_t}{K_{t-1}}}.$$ 

That is, $a_t$ is the proportion of output consumed in period $t$. For the consumption $C_t$ to be feasible, we must have that $0 \leq a_t \leq 1$. If we substitute the above equation for consumption into the Euler equation 10.1 and simplify we obtain the equation

$$a_{t+1} = \beta a_t \sqrt{\frac{K_{t+1}}{K_{t}}}.$$  

(10.5)

If we substitute $\left(1 - a_t\right) 2^{\frac{K_t}{K_{t-1}}}$ for $K_t$ in equation 10.5, assume that $K_{t-1} > 0$, and simplify, we obtain the difference equation

$$a_{t+1} = f(a_t) = \frac{\beta}{2 \frac{1}{a_t}}.$$  

The figure on the previous page indicates that this difference equation has an unstable steady state at $a = 1 - \beta/2$. If $a_0 > a$, then the Euler equation and feasibility force consumption to evolve in such a way that the program becomes infeasible, since $a_t$ eventually exceeds 1. If $a_0 < a$, then the Euler equation and feasibility force consumption to converge to zero, while remaining forever positive. Such a program is suboptimal, because more utility could be had by stopping the decline in consumption at some time and keeping capital at a constant positive level from then on.

I show that the program generated as above with $a_t = a = 1 - \beta/2$, for all $t$, is optimal. We know that if $a_0 < 1 - \beta/2$, then consumption converges to zero and the program is not optimal. We also know that if $a_0 > 1 - \beta/2$ the program eventually becomes infeasible. Therefore the only program consistent with feasibility and the Euler equation and that is not clearly suboptimal is the one with $a_0 = 1 - \beta/2$, and in this case, $a_t = 1 - \beta/2$, for all $t$. Since an optimal program exists by theorem 9.8 and such a program is feasible and satisfies the Euler equation, the program with $a_t = 1 - \beta/2$, for all $t$, must be optimal.

Notice that if $a_0 < a$, then
\[
\frac{dv(C)}{C} \frac{y_t}{t} = \frac{d \ln(C)}{C} \frac{y_t}{t} = \frac{1}{a} \frac{y_t}{t} = \frac{1}{a}
\]

diverges to infinity as \( t \) goes to infinity, since \( a \) converges to 0. The condition that \( \frac{dv(C)}{C} \frac{y_t}{t} \) be bounded is an example of a transversality condition. We will see that this condition plus satisfying the Euler equation are necessary and sufficient for the optimality of feasible programs in problems like this example.

The possibility that a feasible infinite program can satisfy the Euler equation and yet not be optimal is known in growth theory as the Hahn problem after the economist, Frank Hahn, who first pointed it out. The suboptimal infinite feasible paths in the example that satisfy the Euler equation are said to display an over accumulation of capital.

I now turn to the proof of theorem 9.8 that an optimal program exists using the sequence approach. This method starts with a sequence of nearly optimal programs and uses the Bolzano-Weierstrass theorem 4.12 to go to the limit as the sense of nearness to optimality goes to zero.

**Theorem** 9.8: Under assumptions 1 - 4, an optimal program exists.

First of all, I state and prove a useful lemma.

**Lemma** 10.2: Let \( f: Y \to \mathbb{R} \) be a continuous and bounded function, where \( Y \) is a subset of \( \mathbb{R}^n \). If \( (\bar{y}_n^r, y_t^n, \ldots, \) for \( n = 1, 2, \ldots \) is a sequence of sequences such that

\[
\lim_{n \to \infty} y_t^n = y_t \in Y,
\]

for \( t = 0, 1, \ldots \), and if \( 0 < \beta < 1 \), then

\[
\lim_{n \to \infty} \sum_{t=0}^{\infty} \beta^t f(y_t^n) = \sum_{t=0}^{\infty} \beta^t f(y_t).
\]

**Proof:** Since \( f \) is continuous and \( \lim_{n \to \infty} y_t^n = y_t \), for all \( t \), it follows that \( \lim_{n \to \infty} f(y_t^n) = f(y_t) \), for all \( t \). Since \( f \) is bounded, there is a positive number \( b \) such that \( |f(y)| \leq b \), for all \( y \in Y \). Therefore

\[
\left| \sum_{t=T+1}^{\infty} \beta^t f(y_t) \right| \leq \sum_{t=T+1}^{\infty} \beta^t |f(y_t)| \leq b \sum_{t=T+1}^{\infty} \beta^t = \frac{\beta^{T+1} b}{1 - \beta},
\]

for any sequence \( y_0, y_1, \ldots \) in \( Y \) and for any positive integer \( T \). Let \( \varepsilon \) be a positive number, let \( T \)
be so large that \( \frac{\beta^{t+1}b}{1 - \beta} < \frac{\epsilon}{3} \) and let \( N \) be so large that if \( n \geq N \), then \( |f(y^n) - f(y^t)| < \frac{\epsilon}{3(T+1)} \),

for \( t = 0, 1, \ldots, T \). If \( n \geq N \), then

\[
\left| \sum_{t=0}^{\infty} \beta^t f(y^n) - \sum_{t=0}^{\infty} \beta^t f(y^t) \right|
\leq \left| \sum_{t=0}^{T} \beta^t f(y^n) - \sum_{t=0}^{T} \beta^t f(y^t) \right| + \left| \sum_{t=T+1}^{\infty} \beta^t f(y^n) - \sum_{t=0}^{T} \beta^t f(y^t) \right| + \left| \sum_{t=0}^{T} \beta^t f(y^t) - \sum_{t=0}^{\infty} \beta^t f(y^n) \right|
\]

\[
= \left| \sum_{t=T+1}^{\infty} \beta^t f(y^n) \right| + \sum_{t=0}^{T} \beta^t \left| f(y^n) - f(y^t) \right| + \left| \sum_{t=0}^{T} \beta^t f(y^t) \right|
\]

\[
\leq \left| \sum_{t=T+1}^{\infty} \beta^t f(y^n) \right| + \sum_{t=0}^{T} \beta^t \left| f(y^n) - f(y^t) \right| + \left| \sum_{t=0}^{T} \beta^t f(y^t) \right|
\]

\[
< \frac{\epsilon}{3} + (T+1) \frac{\epsilon}{3(T+1)} + \frac{\epsilon}{3} = \epsilon,
\]

Since \( \epsilon \) is arbitrarily small,

\[
\lim_{n \to \infty} \sum_{t=0}^{\infty} \beta^t f(y^n) = \sum_{t=0}^{\infty} \beta^t f(y^t).
\]

**Proof of theorem 9.8:** Since by assumption 4 the domain of \( u \) is compact and \( u \) is continuous, the number \( b = \max_{(x_0, x_1) \in \mathcal{A}} |u(x_0, x_1)| \) exists. Then

\[
\sum_{t=0}^{\infty} \beta^t u(x, x_{t+1}) \leq \frac{b}{1 - \beta},
\]

for all \((x_1, x_2, \ldots, x_{T+1}) \in \mathcal{F}(x_0)\), so that

\[
v = \sup_{(x_1, x_2, \ldots, x_{T+1}) \in \mathcal{F}(x_0)} \sum_{t=0}^{T} \beta^t u(x, x_{t+1}) < \infty.
\]

Assumption 1 implies that \( \mathcal{F}(x_0) \) is not empty, so that \( v > -\infty \). By the definition of the supremum, for each \( n = 1, 2, \ldots \), there exists \((x_1^n, x_2^n, \ldots, x_{T+1}^n) \in \mathcal{F}(x_0)\) such that

\[
v - 1/n < u(x, x_1^n) + \sum_{t=1}^{T} \beta^t u(x, x_{t+1}^n) \leq v.
\]
Because the sequence \( x^n \) is in \( X \) and \( X \) is compact, the Bolzano-Weierstrass theorem 4.12 implies that there is a subsequence \( (n^k)_{k=1}^{\infty} \) of the sequence \( n \) such that \( x = \lim_{k \to \infty} x^{n^k} \) exists and belongs to \( X \). Since \( x^{n^k} \in G(x_0) \), for all \( k \), and \( G \) has closed graph, it follows that \( x \in G(x_0) \).

Suppose that for each \( i = 1, \ldots, S \), we have defined a sequence \( (n^i)_{k=1}^{\infty} \) such that for \( i = 1, \ldots, S - 1 \), \( (n^{i+1})_{k=1}^{\infty} \) is a subsequence of \( (n^i)_{k=1}^{\infty} \) and, for \( i = 1, \ldots, S \) and \( t = 1, \ldots, i \), \( x = \lim_{k \to \infty} x^{n^k} \) exists and belongs to \( X \). Suppose furthermore that \( x \in G(x_0) \) and \( x \in G(x_t) \), for \( t = 1, \ldots, S - 1 \). Because \( \left( x^{n^k}_{S+1, k=1} \right) \) is a sequence in the compact set \( X \), the Bolzano-Weierstrass theorem 4.12 implies that there exists a subsequence \( (n^{S+1})_{k=1}^{\infty} \) of \( (n^S)_{k=1}^{\infty} \) such that \( x = \lim_{k \to \infty} x^{n^k} \) exists and belongs to \( X \). Because \( \left( x^{n^k}_{S+1, k=1} \right) \) is a subsequence of \( \left( x^{n^k}_{S, k=1} \right) \), for \( t = 1, \ldots, S \) and because \( x = \lim_{k \to \infty} x^{n^k} \), for all \( t = 1, \ldots, S \), it follows that \( x = \lim_{k \to \infty} x^{n^k} \), for \( t = 1, \ldots, S \). Since \( x^{n^k}_{S+1} \in G(x^{n^k}_{S}) \), for all \( k \), and \( G \) has closed graph, it follows that \( x \in G(x_0) \).

It has been shown by induction on \( S \) that there exists a sequence of sequences, \( (n^1)_{k=1}^{\infty}, (n^2)_{k=1}^{\infty}, \ldots) \) such that, for each \( S = 1, 2, \ldots \), \( (n^{S+1})_{k=1}^{\infty} \) is a subsequence of \( (n^S)_{k=1}^{\infty} \) and for each \( t = 1, \ldots, S \), \( x = \lim_{k \to \infty} x^{n^k} \) exists and belongs to \( X \). Furthermore \( x \in G(x_0) \) and \( x \in G(x_t) \), for \( t = 1, 2, \ldots \). It follows that \( (x_0, x_1, x_2, \ldots) \in \overline{G}(x_0) \). Because each of \( (x_0, x_1), (x_1, x_2), (x_2, x_3), \ldots \) belongs to the graph of \( G \) and hence to \( A \), it follows that

\[
\sum_{i=1}^t \beta^i u(x_i, x_{i+1})
\]

is well-defined and finite.

I show that the program \( (x_0, x_2, \ldots) \) is optimal given the initial state \( x_0 \). Let \( n = n^k \), for \( k = 1, 2, \ldots \). The definition \( n = n^k \) may perhaps be better understood by considering the following array. The first row represents the sequence \( n \). The second row represents the subsequence \( n^1 \). The second row represents the subsequence \( n^2 \). The third row represents the subsequence \( n^3 \). The entries with an overstrike are the entries corresponding to \( n = n^k \), for \( k = 1, 2, 3 \).
For each positive integer $K$, $n^K_1, n^K_2, \ldots$, is a subsequence of $n^K_{K}, n^K_{K+1}, \ldots$. It follows that if $t = 0, 1, \ldots, K$, then $\lim_{k \to \infty} x^K_t = x$. Since $K$ is an arbitrary positive integer, it follows that $\lim_{k \to \infty} x^K_t = x$, for $t = 0, 1, \ldots$.

By the definition of $(x^K_1, x^K_2, \ldots)$,

$$v - 1/n^S_k \leq u(x^S_0, x^K_1) + \sum_{t=1}^\infty \beta^t u(x^K_t, x^K_{t+1}) \leq v$$

for all $S$ and $k$. Since $n^S_k = n^K_k$, for all $k$, we know that

$$v - 1/n^S_k \leq u(x^S_0, x^K_1) + \sum_{t=1}^\infty \beta^t u(x^K_t, x^K_{t+1}) \leq v,$$

for all $k$. Notice that $n^S_k \geq k$. Let $\varepsilon$ be a small positive number and choose $K$ so that $1/K < \varepsilon/2$, so that $1/n^S_k < \varepsilon/2$. By lemma 10.2, we may assume that $K$ is so large that

$$\left| u(x^S_0, x^K_1) + \sum_{t=1}^\infty \beta^t u(x^K_t, x^K_{t+1}) - u(x^S_0, x) - \sum_{t=1}^\infty \beta^t u(x, x_{t+1}) \right| < \frac{\varepsilon}{2}.$$  

Since

$$v - \frac{\varepsilon}{2} < u(x^S_0, x^K_1) + \sum_{t=1}^\infty \beta^t u(x^K_t, x^K_{t+1}) \leq v,$$

it now follows that

$$v - \varepsilon < u(x^S_0, x) + \sum_{t=1}^\infty \beta^t u(x, x_{t+1}) \leq v.$$

Since $\varepsilon$ is arbitrarily small,
\[ v = u(x_0, x_1) + \sum_{t=1}^{\infty} \beta^t u(x_t, x_{t+1}) \]

and so \((x_0, x_1, \ldots)\) is optimal given the initial state \(x_0\). \[ \qed \]

The type of argument just used to prove theorem 9.8 is known as a Cantor diagonalization argument. The argument is so called because we had a sequence of sequences, \(((n_1^n)^n)_{k=1}^{\infty}, (n_2^n)^n, \ldots\) such that, for each \(S = 1, 2, \ldots\), \((n_1^n)^n\) is a subsequence of \((n_2^n)^n\) and we then used the diagonal subsequence \(n_k^n\), for \(k = 1, 2, \ldots\).

**Definition:** The value function, \(V: X \rightarrow \mathbb{R}\) is defined by the equation

\[ V(x_0) = \max \sum_{(x_t, x_{t+1}) \in \mathcal{A}(x_0)}^{\infty} \beta^t u(x_t, x_{t+1}), \]

for all \(x_0 \in X\).

The next lemma makes rigorous the intuition underlying the Bellman equation 9.8.

**Lemma 10.3:** Let \(f: Y \times Z \rightarrow \mathbb{R}\), where \(Y\) and \(Z\) are sets. If \(\max_{(y, z) \in Y \times Z} f(y, z)\) exists, then

\[ \max_{y \in Y} \left( \sup_{z \in Z} f(y, z) \right) \]

is well-defined and

\[ \max_{y \in Y} \left( \sup_{z \in Z} f(y, z) \right) = \max_{(y, z) \in Y \times Z} f(y, z). \]

Furthermore if \((y, z) \in Y \times Z\) is such that \(f(y, z) = \max_{(y, z) \in Y \times Z} f(y, z)\), then

\[ f(y, z) = \max_{z \in Z} f(y, z) = \max_{y \in Y} \left( \sup_{z \in Z} f(y, z) \right). \]

Conversely if

\[ f(y, z) = \max_{y \in Y} \left( \sup_{z \in Z} f(y, z) \right), \]

for some \((y, z) \in Y \times Z\), then

\[ \max_{(y, z) \in Y \times Z} f(y, z) = f(y, z). \]
Proof: Suppose that \( \max_{(y, z) \in Y \times Z} f(y, z) \) exists. Since

\[
f(y, z) \leq \max_{(y, z) \in Y \times Z} f(y, z)
\]

for all \((y, z) \in Y \times Z\), it follows that

\[
\sup_{z \in Z} f(y, z) \leq \max_{(y, z) \in Y \times Z} f(y, z),
\]

for all \(y \in Y\). Hence

\[
\sup_{y \in Y} \left( \sup_{z \in Z} f(y, z) \right) \leq \max_{(y, z) \in Y \times Z} f(y, z).
\]

Therefore, if \(y\) and \(z\) are such that \(f(y, z) = \max_{(y, z) \in Y \times Z} f(y, z)\), then

\[
f(y, z) \leq \sup_{y \in Y} \left( \sup_{z \in Z} f(y, z) \right) \leq \max_{y \in Y} \left( \sup_{z \in Z} f(y, z) \right) = \max_{(y, z) \in Y \times Z} f(y, z),
\]

and hence

\[
f(y, z) = \max_{y \in Y} \left( \sup_{z \in Z} f(y, z) \right) = \max_{(y, z) \in Y \times Z} f(y, z),
\]

as was to be proved.

Suppose that for some \((y, z) \in Y \times Z\),

\[
f(y, z) = \max_{y \in Y} \left( \sup_{z \in Z} f(y, z) \right).
\]

Since

\[
f(y, z) \leq \sup_{y \in Y} \left( \sup_{z \in Z} f(y, z) \right)
\]

for all \((y, z) \in Y \times Z\), it follows that

\[
\sup_{(y, z) \in Y \times Z} f(y, z) \leq \max_{y \in Y} \left( \sup_{z \in Z} f(y, z) \right) = f(y, z).
\]

Therefore
\[ f(y, z) = \max_{(y, z) \in Y \times Z} f(y, z) \]

and so

\[
\max_{(y, z) \in Y \times Z} f(y, z) = f(\bar{y}, \bar{z}) = \max\{\sup_{y \in Y} \sup_{z \in Z} f(y, z)\}.
\]

This lemma has the following important corollary concerning the value function \( V \).

**Corollary 10.4:** Suppose that assumptions 1 - 4 apply. If \( x_0 \in X \), then

\[
V(x_0) = \max_{x_1 \in G(x_0)} \left[ u(x_0, x_1) + \beta V(x_1) \right].
\]

Furthermore if \( (x_1, x_2, \ldots) \in F(x_0) \) is optimal, then

\[
V(x_0) = u(x_0, x_1) + \beta V(x_1),
\]

and, for \( t = 1, 2, \ldots \), \( (x_i, x_{i+1}, \ldots) \in F(x_i) \) is optimal, so that

\[
V(x_i) = \sum_{s=1}^{t} \beta^s u(x_s, x_{s+1}).
\]

In addition,

\[
V(x_1) = u(x_0, x_1) + \beta V(x_{1+1}) = \max_{x_{1+1} \in G(x_1)} \left[ u(x_0, x_{1+1}) + \beta V(x_{1+1}) \right]
\]

**Proof:** Let \( b = \max_{(x, y) \in A} |u(x, y)| \), where the maximum exists because \( A \) is compact and non-empty and \( u \) is continuous. I apply the first half of lemma 10.3. Let the \( Y \) of the lemma equal the \( X \) of the dynamic programming model. Let \( Z \) be the set of all feasible programs \( (x_1, x_2, \ldots) \in F(x) \), for any initial state \( x_0 \in X \). Fix \( x_0 \) and let

\[
f(x_1; x_2, x_3, \ldots) = \begin{cases} 
\sum_{i=0}^{\infty} \beta^i u(x_i, x_{i+1}), & \text{if } x_i \in G(x_0), \\
-\frac{2b}{1-\beta}, & \text{otherwise}.
\end{cases}
\]

The number \(-2b/(1-\beta)\) is so large in the negative direction that \( f \) is maximized only by feasible programs. That is,
\[
\max_{(x_1, x_2, x_3, \ldots)} f(x_1, x_2, x_3, \ldots) = \max_{(x_1, x_2, x_3, \ldots) \in \mathcal{Y}^2 \times \mathcal{Z}} f(x_1, x_2, x_3, \ldots). \tag{10.6}
\]

By theorem 9.8, the maximum on the right-hand side of equation 10.6 exists. It equals \( V(x_0) \).

Similarly
\[
\max_{(x_2, x_3, \ldots) \in \mathcal{Z}} f(x_1, x_2, x_3, \ldots) = \max_{(x_2, x_3, \ldots) \in \mathcal{F}(x_1)} f(x_1, x_2, x_3, \ldots), \tag{10.7}
\]

because if \( x_2 \) does not belong to \( \mathcal{G}(x_1) \), then \( f(x_1, x_2, x_3, \ldots) = \frac{-2b}{1 - \beta} \), which is less than \( \sum_{t=0}^{\infty} \beta^t u(x_t, x_1) \), for any choice of \((x_2, x_3, \ldots)\) in \( \mathcal{F}(x_1) \). The maximum on the right-hand side of equation 10.7 exists by theorem 9.8 and equals \( u(x_1, x_0) + \beta V(x_1) \), if \( x_1 \in \mathcal{G}(x_0) \) and equals \( -2b/(1 - \beta) \) otherwise. By the same kind of reasoning,
\[
\max_{x_1 \in \mathcal{Y}} \left( \max_{(x_2, x_3, \ldots) \in \mathcal{Z}} f(x_1, x_2, x_3, \ldots) \right) = \max_{x_1 \in \mathcal{G}(x_0)} \left( \max_{(x_2, x_3, \ldots) \in \mathcal{F}(x_1)} f(x_1, x_2, x_3, \ldots) \right) \tag{10.8}
\]

It now follows that
\[
V(x_0) = \max_{(x_1, x_2, x_3, \ldots) \in \mathcal{Y} \times \mathcal{Z}} f(x_1, x_2, x_3, \ldots) = \max_{x_1 \in \mathcal{Y}} \left( \sup_{(x_2, x_3, \ldots) \in \mathcal{Z}} f(x_1, x_2, x_3, \ldots) \right) \tag{10.9}
\]

\[
= \max_{x_1 \in \mathcal{Y}} \left( \max_{(x_2, x_3, \ldots) \in \mathcal{Z}} f(x_1, x_2, x_3, \ldots) \right) = \max_{x_1 \in \mathcal{G}(x_0)} \left( \max_{(x_2, x_3, \ldots) \in \mathcal{F}(x_1)} f(x_1, x_2, x_3, \ldots) \right) \tag{10.9}
\]

\[
= \max_{x_1 \in \mathcal{G}(x_0)} \left[ u(x_0, x_1) + \beta V(x_1) \right],
\]

where the first and last equations follow from the definition of \( f \) and \( V \) and theorem 9.8, the third equation follows from theorem 9.8, the second equation follows from the first half of lemma 10.3, and the fourth equation is equation 10.8. If \((x_1, x_2, \ldots) \in \mathcal{F}(x_0)\) is such that
\[
f(x_1, x_2, x_3, \ldots) = \max_{(x_1, x_2, x_3, \ldots) \in \mathcal{Y} \times \mathcal{Z}} f(x_1, x_2, x_3, \ldots), \tag{10.10}
\]

then
\[
f(x_1, x_2, x_3, \ldots) = \max_{(x_2, x_3, \ldots) \in \mathcal{Z}} f(x_1, x_2, x_3, \ldots) = \max_{x_1 \in \mathcal{Y}} \left( \sup_{(x_2, x_3, \ldots) \in \mathcal{Z}} f(x_1, x_2, x_3, \ldots) \right) \tag{10.11}
\]

\[
= \max_{x_1 \in \mathcal{Y}} \left( \max_{(x_2, x_3, \ldots) \in \mathcal{Z}} f(x_1, x_2, x_3, \ldots) \right) = \max_{x_1 \in \mathcal{G}(x_0)} \left( \max_{(x_2, x_3, \ldots) \in \mathcal{F}(x_1)} f(x_1, x_2, x_3, \ldots) \right).
\]
where the first and second equations follow from the first half of lemma 10.3, the third equation follows from theorem 9.8, and the last equation follows from equation 10.9. Since 

\[ x, x, .... \in \mathcal{F}(x_0) \], it follows from the definition of the function f that

\[ f(x_0, x, x, ....) = u(x_0, x) + \sum_{l=1}^{\infty} \beta^l u(x_l, x_{l+1}). \]

Therefore equations 10.9 - 10.11 imply that

\[ V(x_0) = u(x_0, x) + \sum_{l=1}^{\infty} \beta^l u(x_l, x_{l+1}) = u(x_0, x) + \beta V(x_1) \]

\[ = \max_{x_1 \in G(x_0)} [u(x_0, x_1) + \beta V(x_1)]. \]

If we apply the same argument with initial time 1 and then initial time 2 and so on, successively, we find that for \( t = 1, 2, .... \), \( x, x, .... \in \mathcal{F}(x) \) is optimal,

\[ V(x_t) = u(x_t, x_{t+1}) + \beta V(x_{t+1}) \]

and

\[ V(x_t) = u(x_t, x_{t+1}) + \beta V(x_{t+1}) = \max_{x_{t+1} \in G(x_t)} [u(x_t, x_{t+1}) + \beta V(x_{t+1})]. \]

It has been shown that the maximum appearing in the Bellman equation exists without knowing whether the value function V is continuous and hence whether the function \( u(x_0, x_1) + \beta V(x_1) \) achieves a maximum over any compact set in its domain. In an appendix to the next lecture, I will show by means of an example that V may not be continuous. In the same lecture, I will show that V is continuous if G is lower semicontinuous.