# Lecture #11

I next introduce a monotonicity assumption that makes it possible to simplify condition 10.8.

Assumption 8: The function u(x, y) is non-decreasing with respect to the components of x, and $X \subset R^N_+$.

Lemma 11.1: Suppose that assumption 8 applies and that $(x, y) \in \text{int } A$. If $(p, q) \in R^N \times R^N$ is a subgradient of u at (x, y), then $p_n \geq 0$, for all n.

Proof: Because $(x, y) \in \text{int } A$, if the positive number $\varepsilon$ is sufficiently small, $(x + \varepsilon e_n, y) \in A$, where $e_n$ is the $n^{th}$ standard basis vector of $R^N$. Because (p, q) is a subgradient of u at (x, y),

$$u(x + \varepsilon e_n, y) \leq u(x, y) + p.(x + \varepsilon e_n - x) = u(x, y) + \varepsilon p_n.$$

Therefore

$$p_n \geq \varepsilon^{-1}[u(x + \varepsilon e_n, y) - u(x, y)] \geq 0,$$

where the last inequality holds, because by assumption 8 u is non-decreasing with respect to the components of x. ■

Suppose that assumptions 7 and 8 apply and that the program $(\underline{x}_1, \underline{x}_2, \ldots) \in \mathfrak{F}(\underline{x}_0)$ is such that $\underline{x}_0 \in \text{int } X$ and $\underline{x}_t \in \text{int } G(\underline{x}_{t-1})$, for t = 1, 2, .... . Because G is lower semicontinuous, lemma 10.8 implies that $(\underline{x}_t, \underline{x}_{t+1})$, for all t, belongs to the interior of the graph of G and so belongs to the interior of A, since A contains the graph of G. Hence the assumptions of lemma 11.1 apply, so that if $(p_0, p_1, \ldots)$ is a sequence of Euler subgradients associated $(\underline{x}_1, \underline{x}_2, \ldots)$, the vectors $p_t$ are non-negative. In addition, the vectors x in X are non-negative by assumption 8. Therefore $-p_t.(x - \underline{x}_t) \leq p_t.\underline{x}_t$ and hence the transversality condition of theorem 11.6 can be replaced by the transversality condition

$$\sup_{t = 1, 2, \ldots} p_t.\underline{x}_t < \infty.$$

Theorem 10.11 and a slight modification of the proof of theorem 10.12 imply the following.

Theorem 11.2: Suppose that assumptions 1 - 8 apply. Let $\underline{x}_0 \in$ int X and suppose that $(\underline{x}_1, \underline{x}_2, \ldots) \in \mathfrak{F}(\underline{x}_0)$ is such that $\underline{x}_{t+1} \in \text{int } G(\underline{x}_t)$, for t = 0, 1, .... . Then $(\underline{x}_1, \underline{x}_2, \ldots)$ is optimal if it has an associated Euler sequence, $(p_0, p_1, \ldots)$, of subgradients such that

$$\sup_{t = 1, 2, \ldots} p_t . \underline{x}_t < \infty .$$

I now use differentiability mainly to restate the Euler condition.

Assumption 9: The function $u : A \to R$ is differentiable on int A.

In order to have convenient notation for the derivatives of u, I write u(x, y), where $x \in X$ and $y \in X$. Then $D_x u(x_t, x_{t+1})$ denotes the derivative of u at $(x_t, x_{t+1})$ with respect to $x_t$, and $D_y u(x_t, x_{t+1})$ denotes the derivative of u at $(x_t, x_{t+1})$ with respect to $x_{t+1}$.

It should be clear that if p is a subgradient of a differentiable function $f: C \to R$ at $x \in$ int C, where C is a subset of $R^N$, then p = Df(x). Furthermore if f is concave, then Df(x) is a subgradient of f at x. These two facts imply that derivatives and subgradients are one and the same for functions that are differentiable and concave. I use this idea now in proving that the value function is differentiable.

Theorem 11.3: Suppose that assumptions 1 - 7 and 9 apply. If $\underline{x}_0 \in$ int X and $\underline{x}_1 \in \text{int } G(\underline{x}_0)$, where $(\underline{x}_1, \underline{x}_2, \ldots) \in \mathfrak{F}(\underline{x}_0)$ is optimal, then V is differentiable at $\underline{x}_0$ and $DV(\underline{x}_0) = D_x u(\underline{x}_0, \underline{x}_1)$.

Proof: By proposition 10.7, the value function has a subgradient p at $\underline{x}_0$. By theorem 10.10, p is a subgradient of the function $g(x_0) = u(x_0, \underline{x}_1)$ at $\underline{x}_0$. Because u is differentiable by assumption 9,

$$p = D_x u(\underline{x}_0, \underline{x}_1) . \qquad (11.1)$$

By lemma 10.8, there is a positive number ε such that if $||x_0 - \underline{x}_0|| < \varepsilon$, then $\underline{x}_1 \in G(x_0)$ and so

$$V(x_0) \geq u(x_0, \underline{x}_1) + \sum_{t=1}^{\infty} \beta^t u(\underline{x}_t, \underline{x}_{t+1}) .$$

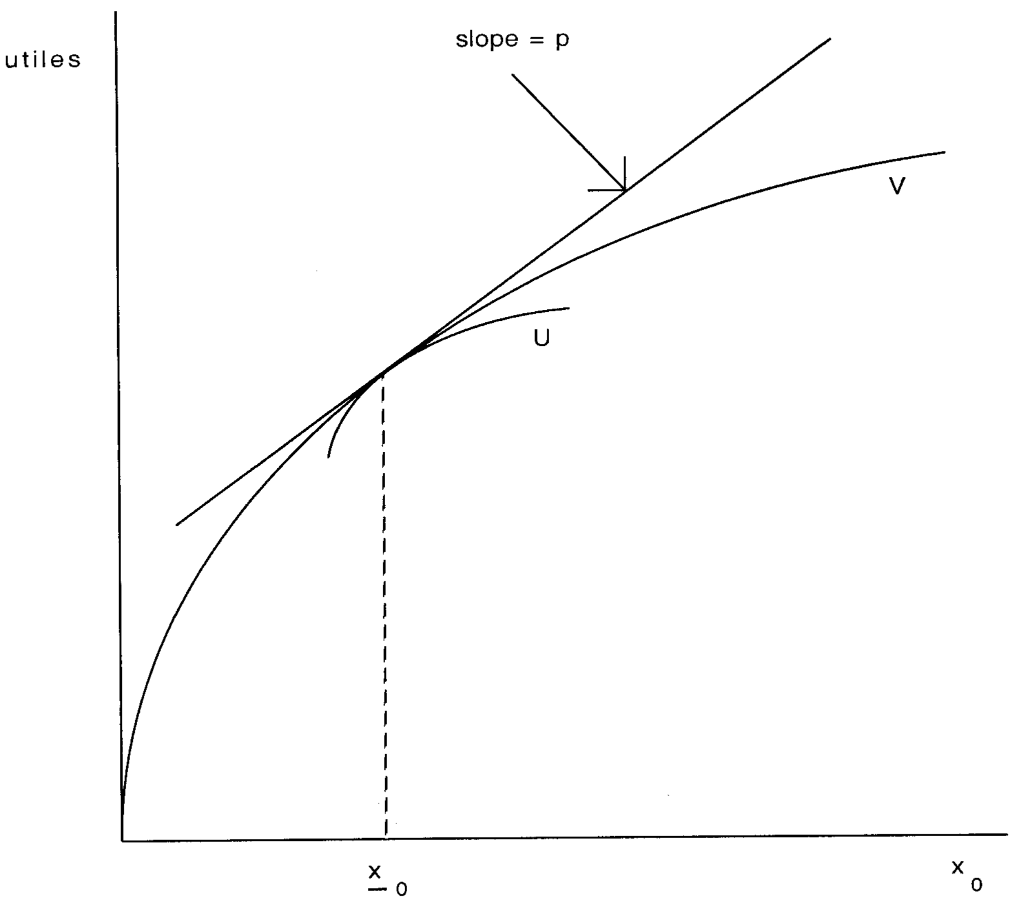
Let $U(x_0) = u(x_0, \underline{x}_1) + \sum_{t=1}^{\infty} \beta^t u(\underline{x}_t, \underline{x}_{t+1})$. Then U is differentiable, $DU(\underline{x}_0) = D_x u(\underline{x}_0, \underline{x}_1)$ = p, and for $x_0$ such that $||x_0 - \underline{x}_0|| < \varepsilon$,

$$U(x_0) \leq V(x_0) \leq V(\underline{x}_0) + p.(x_0 - \underline{x}_0)\,, \qquad (11.2)$$

and furthermore $U(\underline{x}_0) = V(\underline{x}_0)$. Subtracting $V(\underline{x}_0) + p.(x_0 - \underline{x}_0)$ from each member of the inequalities 11.2, we see that

$$u(x_0, \underline{x}_1) - u(\underline{x}_0, \underline{x}_1) - p.(x_0 - \underline{x}_0) = U(x_0) - V(\underline{x}_0) - p.(x_0 - \underline{x}_0)$$

$$\leq V(x_0) - V(\underline{x}_0) - p.(x_0 - \underline{x}_0) \leq 0\,.$$




Since $D_x u(\underline{x}_0, \underline{x}_1) = p$,

$$\lim_{\substack{x_0 \to \underline{x}_0 \\ x_0 \neq \underline{x}_0}} \frac{\left| u(x_0, \underline{x}_1) - u(\underline{x}_0, \underline{x}_1) - p.(x_0 - \underline{x}_0) \right|}{\|x_0 - \underline{x}_0\|} = 0.$$

Hence

$$\lim_{\substack{x_0 \to \underline{x}_0 \\ x_0 \neq \underline{x}_0}} \frac{\left| V(x_0) - V(\underline{x}_0) - p.(x_0 - \underline{x}_0) \right|}{\|x_0 - \underline{x}_0\|} = 0$$

and so $DV(\underline{x}_0) = p$, by the definition of the derivative of V at $\underline{x}_0$. From equation 11.1, it follows that $DV(\underline{x}_0) = D_x u(\underline{x}_0, \underline{x}_1)$. The argument is illustrated in the figure on the previous page. ■

The next theorem is a corollary of theorems 10.11 and 11.3.

Theorem 11.4: Suppose that assumptions 1 - 7 and 9 apply, that $\underline{x}_0 \in$ int X, and that $(\underline{x}_1, \underline{x}_2, \ldots) \in \mathfrak{F}(\underline{x}_0)$ is optimal and is such that $\underline{x}_{t+1} \in$ int $G(\underline{x}_t)$, for t = 0, 1, ..... . For each t, let $p_t = D_x u(\underline{x}_t, \underline{x}_{t+1})$. Then for all t, $p_t$ is a subgradient and is the derivative of V at $\underline{x}_t$ and $(p_t, -\beta p_{t+1}) = Du(\underline{x}_t, \underline{x}_{t+1})$ and is a subgradient of u at $(\underline{x}_t, \underline{x}_{t+1})$, so that $D_y u(\underline{x}_t, \underline{x}_{t+1}) = -\beta D_x u(\underline{x}_{t+1}, \underline{x}_{t+2})$.

The equation $D_y u(\underline{x}_t, \underline{x}_{t+1}) = -\beta D_x u(\underline{x}_{t+1}, \underline{x}_{t+2})$ is known as the Euler equation. The next theorem is a corollary of theorems 11.2 and 11.4.

Theorem 11.5: Suppose that assumptions 1 - 9 apply. Let $\underline{x}_0 \in$ int X and suppose that $(\underline{x}_1, \underline{x}_2, \ldots) \in \mathfrak{F}(\underline{x}_0)$ is such that $\underline{x}_{t+1} \in$ int $G(\underline{x}_t)$, for t = 0, 1, .... . Then $(\underline{x}_1, \underline{x}_2, \ldots)$ is optimal if $D_y u(\underline{x}_t, \underline{x}_{t+1}) = -\beta D_x u(\underline{x}_{t+1}, \underline{x}_{t+2})$, for t = 0, 1, .... and
$\sup_{t = 1, 2, \ldots} D_x u(\underline{x}_t, \underline{x}_{t+1}).\underline{x}_t < \infty$.

The condition $\sup_{t = 1, 2, \ldots} D_x u(\underline{x}_t, \underline{x}_{t+1}).\underline{x}_t < \infty$ is another transversality condition.

Remark: If u is differentiable, then it is not necessary to use convex analysis to derive the Euler equation. Suppose that G is lower semicontinuous and has convex graph and that $(\underline{x}_1, \underline{x}_2, \ldots) \in \mathfrak{F}(\underline{x}_0)$ is optimal and such that $\underline{x}_t \in \text{int } G(\underline{x}_{t-1})$, for t = 1, 2, .... . For each T = 1, 2, .... , there is, by lemma 10.8, a positive number $\varepsilon$ such that if $x_T \in R^N$ and $||x_T - \underline{x}_T|| < \varepsilon$, then $x_T \in G(\underline{x}_{T-1})$ and $\underline{x}_{T+1} \in G(x_T)$. Therefore $u(\underline{x}_{T-1}, x_T)$ and $u(x_T, \underline{x}_{T+1})$ are well-defined if $||x_T - \underline{x}_T|| < \varepsilon$, and so the derivative $D_{x_T}[u(\underline{x}_{T-1}, \underline{x}_T) + \beta u(\underline{x}_T, \underline{x}_{T+1})]$ makes sense. Since $(\underline{x}_1, \underline{x}_2, \ldots)$ is optimal,

$$0 = D_{x_T}\left[\sum_{t=0}^{\infty} \beta^t u(\underline{x}_t, \underline{x}_{t+1})\right] = D_{x_T}[\beta^{T-1}u(\underline{x}_{T-1}, \underline{x}_T) + \beta^T u(\underline{x}_T, \underline{x}_{T+1})$$

$$= \beta^{T-1} D_y u(\underline{x}_{T-1}, \underline{x}_T) + \beta^T D_x u(\underline{x}_T, \underline{x}_{T+1}).$$

The Euler equation $D_y u(\underline{x}_{T-1}, \underline{x}_T) = -\beta D_x u(\underline{x}_T, \underline{x}_{T+1})$ follows immediately.

## Contraction Mapping Theorem

I now turn to the contraction mapping theorem, which gives an approach to proving the existence of an optimal program that is different from that presented in theorem 10.1. The idea is to obtain the value function as a limit of successive approximations. Suppose we guess at a trial value function $V^0: X \to R$. We might use this to calculate another value function, $V^1(x_0) = \max_{x_1 \in G(x_0)} [u(x_0, x_1) + \beta V^0(x_1)]$. We could proceed inductively; given $V^T$, let $V^{T+1}(x_0) = \max_{x_1 \in G(x_0)} [u(x_0, x_1) + \beta V^T(x_1)]$. Then $V^T$ is the same as the function defined by the equation

$$V^T(x_0) = \max_{x_1 \in G(x_0), \ldots, x_T \in G(x_{T-1})} \left[\sum_{t=0}^{T-1} \beta^t u(x_t, x_{t+1}) + \beta^T V^0(x_T)\right].$$

Since $\beta^T$ converges to zero as T goes to infinity, the choice of $V^0$ should not matter and $V^T$ should converge to the true value function. In order to develop this intuition, I need some mathematical preliminaries.

If X is a non-empty compact subset of $R^N$, let C(X) be the set of real-valued continuous functions on X. That is, $C(X) = \{f: X \to R \mid f \text{ is continuous}\}$. If $f \in C(X)$, the maximum norm of f is $\|f\| = \max_{x \in X} |f(x)|$. Because X is compact and f is continuous, $\|f\|$ exists and is finite. Say that

$\lim_{n\to\infty} f_n = f$, if $\lim_{n\to\infty} \|f_n - f\| = 0$.

Lemma 11.6: If $f_n$ is a sequence in C(X) and $\lim_{n\to\infty} f_n = f$, where $f: X \to R$, then f is continuous and so belongs to C(X).

Proof: Let $x_n$ be a sequence in X converging to $x \in X$. I must show that $\lim_{n\to\infty} f(x_n) = f(x)$. Let $\varepsilon$ be a positive number. Choose the positive integer K so large that $\|f_K - f\| < \varepsilon/3$. Since $f_K$ is continuous, there exists a positive integer N such that $|f_K(x_n) - f_K(x)| < \varepsilon/3$, if $n \geq N$. If $n \geq N$,

$$|f(x_n) - f(x)| \leq |f(x_n) - f_K(x_n)| + |f_K(x_n) - f_K(x)| + |f_K(x) - f(x)|$$

$$< \|f - f_K\| + \frac{\varepsilon}{3} + \|f_K - f\| < \frac{\varepsilon}{3} + \frac{\varepsilon}{3} + \frac{\varepsilon}{3} = \varepsilon.$$

Since $\varepsilon$ may be arbitrarily small, $\lim_{n\to\infty} f(x_n) = f(x)$. ■

Lemma 11.7: C(X) is complete with respect to the norm $\|.\|$. That is, if $f_n$ is a sequence in C(X) that is Cauchy with respect to $\|.\|$, then there is an $f \in C(X)$ such that $\lim_{n\to\infty} f_n = f$.

Proof: The sequence $f_n$ is Cauchy if

$$\lim_{N\to\infty} \sup_{\substack{n \geq N, \\ m \geq N}} \|f_n - f_m\| = 0.$$

If $f_n$ is Cauchy, then for each $x \in X$, $|f_n(x) - f_m(x)| \leq \|f_n - f_m\|$, so that the sequence $f_n(x)$ is a Cauchy sequence of numbers. Since the real numbers are complete, there exists a number f(x) such that $\lim_{n\to\infty} f_n(x) = f(x)$. Therefore there is a function $f: X \to R$ such that $\lim_{n\to\infty} f_n(x) = f(x)$, for all x. I show that $\lim_{n\to\infty} \|f_n - f\| = 0$. Let $\varepsilon$ be a positive number. Since $f_n$ is Cauchy, there exists a positive integer N such that $\|f_n - f_m\| < \varepsilon/2$, if $n \geq N$ and $m \geq N$. If $x \in X$, there is a positive integer k, depending on x, such that $k \geq N$ and $|f_k(x) - f(x)| < \varepsilon/2$. If $n \geq N$, then

$$|f_n(x) - f(x)| \leq |f_n(x) - f_k(x)| + |f_k(x) - f(x)| \leq \|f_n - f_k\| + |f_k(x) - f(x)|$$

$$< \varepsilon/2 + \varepsilon/2 = \varepsilon.$$

Since x is an arbitrary point in X, $\|f_n - f\| < \varepsilon$, if $n \geq N$. Since $\varepsilon$ may be arbitrarily small, $\lim_{n \to \infty} \|f_n - f\| = 0$. By the previous lemma, $f \in C(X)$. Therefore C(X) is complete. ■

Definition: A function $Q: C(X) \to C(X)$ is a contraction if for some $\beta$ such that $0 < \beta < 1$, $\|Q(f) - Q(g)\| \leq \beta\|f - g\|$, for all f and g in C(X). The number $\beta$ is called the modulus of Q.

Definition: If $Q: C(X) \to C(X)$, a fixed point of Q is an $f \in C(X)$ such that $Q(f) = f$.

Contraction Mapping Theorem 11.8: Every function $Q: C(X) \to C(X)$ that is a contraction has a unique fixed point.

Proof: Let $\beta$ be the modulus of Q, where $0 < \beta < 1$. I show that the sequence f, Q(f), $Q(Q(f)) = Q^2(f)$, $Q(Q(Q(f))) = Q^3(f)$, .... in C(X) is Cauchy with respect to $\|.\|$.

$$\|Q^{n+1}(f) - Q^n(f)\| = \|Q(Q^n(f)) - Q(Q^{n-1}(f))\| \leq \beta\|Q^n(f) - Q^{n-1}(f)\| \leq ....$$

$$\leq \beta^n\|Q(f) - Q^0(f)\| = \beta^n\|Q(f) - f\|,$$

where the last equation holds because $Q^0(f) = f$. Therefore if $n > m \geq N$, then

$$\|Q^n(f) - Q^m(f)\| \leq \|Q^n(f) - Q^{n-1}(f)\| + \|Q^{n-1}(f) - Q^{n-2}(f)\| + .... + \|Q^{m+1}(f) - Q^m(f)\|$$

$$\leq (\beta^{n-1} + \beta^{n-2} + .... + \beta^m)\|Q(f) - f\| \leq (\beta^{n-1} + \beta^{n-2} + .... + \beta^N)\|Q(f) - f\|$$

$$\leq (.... + \beta^{N+1} + \beta^N)\|Q(f) - f\| = \frac{\beta^N}{1 - \beta}\|Q(f) - f\|.$$

Since $\lim_{N \to \infty} \frac{\beta^N}{1 - \beta}\|Q(f) - f\| = 0$, the sequence $Q^n(f)$ is Cauchy with respect to $\|.\|$, so that the previous lemma implies that there exists a $g \in C(X)$ such that $\lim_{n \to \infty} \|Q^n(f) - g\| = 0$. Since $\|Q^{n+1}(f) - Q(g)\| \leq \beta\|Q^n(f) - g\|$, it follows that $\lim_{n \to \infty} \|Q^{n+1}(f) - Q(g)\| = 0$, which is the same as saying that $\lim_{n \to \infty} \|Q^n(f) - Q(g)\| = 0$. Since $\lim_{n \to \infty} Q^n(f) = g$ and $\lim_{n \to \infty} Q^n(f) = Q(g)$, it follows that $Q(g) = g$ and so g is a fixed point of Q.

I show that the fixed point is unique. Suppose that g and h are fixed points of Q. Because Q is a contraction with modulus $\beta$,

$$\|g - h\| = \|Q(g) - Q(h)\| \leq \beta\|g - h\|.$$

Since $\beta < 1$, this inequality is possible only if $\|g - h\| = 0$, that is, if $g = h$. ■

The next task is to apply the contraction mapping theorem to the dynamic programming model. Let X be as in that model and if $f \in C(X)$, let $Q(f) : X \to R$ be defined by the equation

$$Q(f)(x_0) = \max_{x_1 \in G(x_0)} [u(x_0, x_1) + \beta f(x_1)]. \tag{11.3}$$

Since $G(x_0)$ is compact and non-empty and u and f are continuous, the maximum exists and Q(f) is well-defined. I check that Q(f) maps into C(X).

Lemma 11.9: If assumptions 1 - 4, and 7 apply and $f \in C(X)$, then Q(f) belongs to C(X), where Q(f) is defined by equation 11.3.

Proof: Q(f) satisfies all the assumptions of Berge's theorem 5.1, because G is an upper and lower semicontinuous correspondence, G(x) is non-empty, for all x, X is compact and non-empty, and u is continuous. Therefore Berge's theorem implies that Q(f) is continuous. ■

Lemma 11.10: If assumptions 1 - 4, and 7 apply, then $Q: C(X) \to C(X)$ is a contraction with respect to $\|.\|$ of modulus β, where Q(f) is defined by equation 11.3.

Proof: Let f and g belong to C(X) and let $x_0 \in X$. For some $x_1^f \in G(x_0)$,

$$Q(f)(x_0) = u(x_0, x_1^f) + \beta f(x_1^f),$$

and for some $x_1^g \in G(x_0)$,

$$Q(g)(x_0) = u(x_0, x_1^g) + \beta g(x_1^g).$$

By the definition of $Q(f)(x_0)$ as a maximum,

$$Q(f)(x_0) \geq u(x_0, x_1^g) + \beta f(x_1^g)$$

$$\geq u(x_0, x_1^g) + \beta g(x_1^g) - \beta\|f - g\| = Q(g)(x_0) - \beta\|f - g\|.$$

By the symmetric argument,

$$Q(g)(x_0) \geq Q(f)(x_0) - \beta\|f - g\|.$$

Therefore

$$\left|Q(f)(x_0) - Q(g)(x_0)\right| \leq \beta\|f - g\|,$$

for all $x_0 \in X$, and so

$$\|Q(f) - Q(g)\| = \max_{x_0 \in X} \left|Q(f)(x_0) - Q(g)(x_0)\right| \leq \beta\|f - g\|.$$

That is, Q is a contraction with modulus $\beta$. ∎

It is now possible to see the connection between the contraction mapping theorem and the existence of an optimal program.

<u>Theorem</u> 11.11: If assumptions 1 - 4, and 7 apply, then for any $x_0 \in X$, there exists an optimal program, $(\underline{x}_1, \underline{x}_2, .....) \in \mathfrak{F}(x_0)$, and the value function $V: X \to R$ is continuous.

<u>Proof</u>: Let $Q: C(X) \to C(X)$ be defined by equation 11.3. By lemma 11.9, Q is well-defined. By lemma 11.10, Q is a contraction. By the contraction mapping theorem 11.8, Q has a unique fixed point F. Because F is a continuous function defined on a compact set, it is bounded. Therefore by theorem 10.5, F equals the value function V. Because $V \in C(X)$, V is continuous. Because the value function exists, an optimal program exists as well. ∎

The next example shows that V may not be continuous if we do not assume that G is lower semicontinuous, even if G is upper semicontinuous and has convex graph and if u is concave.

<u>Example</u>: (Discontinuous value function) Let

$$X = \{(y_1, y_2) \in R^2 \mid (y_1 - 3)^2 + (y_2 - 3)^2 \leq 2\}.$$

As a step in the definition of G, let

$$\Gamma = \{(y_1, y_2, 2, 2) \mid (y_1, y_2) \in X\} \cup \{(4, 4, c, c) \mid 2 \leq c \leq 4\}.$$

Let the graph of G be the convex hull of $\Gamma$, which is the set of all convex combinations of points in $\Gamma$. The graph of G is convex and compact. The correspondence G is defined by the equation

$$G(y_1, y_2) = \{(y_3, y_4) \mid (y_1, y_2, y_3, y_4) \in \text{graph of } G\}.$$

Let $A = X \times X$ and let $u: X \times X \to R$ be defined by the equation

$$u(x_1, x_2) = x_{21} + x_{22}$$

and let $\beta \in (0, 1)$. The set G(x) is not hard to visualize. Let C be the circle that is the boundary of X. That is,

$$C = \{(y_1, y_2) \in R^2 \mid (y_1 - 3)^2 + (y_2 - 3)^2 = 2\},$$

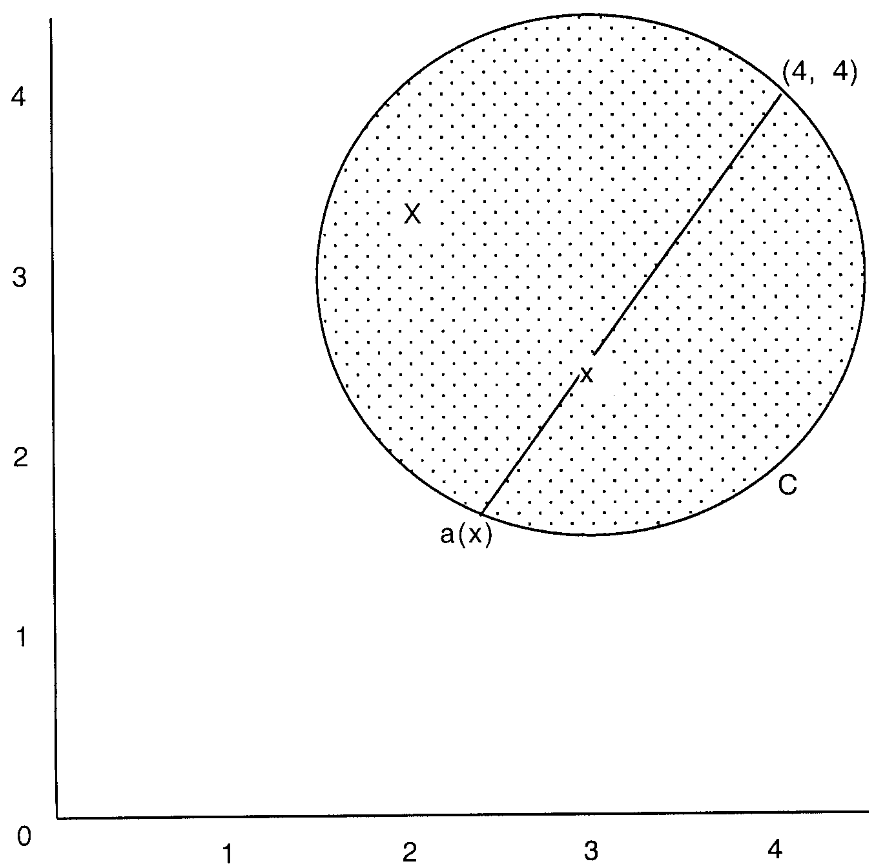
so that $(4, 4) \in C$. If $x \in X$ and $x \neq (4, 4)$, there is a unique line segment going from (4, 4) through x. Let a(x) be the intersection of C with this line segment. Then x is a convex combination of a(x) and (4, 4). Let $\alpha(x)$ be defined by the equation

$$x = \alpha(x)a(x) + (1 - \alpha(x))(4, 4).$$

Let $\alpha(4, 4) = 0$. Notice that $\alpha(x) = 1$, if $x \in C$ and $x \neq (4, 4)$. Then

$$G(x) = \{(r, r) \mid 2 \leq r \leq 4 - 2\alpha(x)\}.$$

G is upper semicontinuous, because its graph is a closed set. G(x) is not empty, for all x. The sets X and A are compact and u is continuous. The sets A and X and the graph of G are convex, X has non-empty interior, and u is concave, so that the example satisfies assumptions 1 - 6. I show that the correspondence G is not lower semicontinuous at the point (4, 4). Let $x^n$ be a sequence in C such that $x^n \neq (4, 4)$, for all n, and $\lim_{n \to \infty} x^n = (4, 4)$. Then $G(x^n) = \{(2, 2)\}$, for all n, and $G((4, 4)) = \{(r, r) \mid 2 \leq r \leq 4\}$. Hence there is no sequence $y^n$ such that $y^n \in G(x^n)$, for all n, and $\lim_{n \to \infty} y^n = (4, 4)$, and so G is not lower semicontinuous at (4, 4) and the example does not satisfy assumption 7. The next figure illustrates the definition of a(x).



If $x_0 = (4, 4)$, then it is optimal to choose $x_1 = (4, 4)$ and enjoy utility 8 in period 0. The optimal choice is then repeated indefinitely and a utility of 8 is enjoyed in every period. Therefore $V(4, 4) = 8/(1 - \beta)$. If $x_0$ is any point in the circle C other than (4, 4), then the only choice possible is (2, 2), which is in C and not equal to (4, 4), and so (2, 2) must be chosen in period 1 and in every period thereafter. Therefore $V(x_0) = 4/(1 - \beta)$, and so the value function is not continuous at (4, 4). This ends the discussion of the example.

A typical application of the contraction mapping theorem is the proof that the value function is non-decreasing. (Of course, this assertion may be proved using the sequence approach as well.) For this result, we need an additional assumption.

<u>Assumption</u> 10: If $\underline{x}_0$ and $\bar{x}_0$ belong to X and $\underline{x}_{0,n} \leq \bar{x}_{0,n}$, for all n, then $G(\underline{x}_0) \subset G(\bar{x}_0)$.

<u>Lemma</u> 11.12: If assumptions 1 - 4, 7, 8 and 10 apply, then the value function V is non-decreasing.

<u>Proof</u>: Let $Q: C(X) \to C(X)$ be defined by equation 11.3. I show that if $f \in C(X)$, then $Q(f)(x_0)$ is a non-decreasing function of $x_0$. Let $\underline{x}_0$ and $\bar{x}_0$ be points in X such that $\underline{x}_0 \leq \bar{x}_0$. If $\underline{x}_1 \in G(\underline{x}_0)$ is such that

$$Q(f)(\underline{x}_0) = u(\underline{x}_0, \underline{x}_1) + \beta f(\underline{x}_1),$$

then

$$Q(f)(\underline{x}_0) = u(\underline{x}_0, \underline{x}_1) + \beta f(\underline{x}_1) = \max_{x_1 \in G(\underline{x}_0)} [u(\underline{x}_0, x_1) + \beta f(x_1)]$$

$$\leq \max_{x_1 \in G(\bar{x}_0)} [u(\underline{x}_0, x_1) + \beta f(x_1)] \leq \max_{x_1 \in G(\bar{x}_0)} [u(\bar{x}_0, x_1) + \beta f(x_1)] = Q(f)(\bar{x}_0),$$

where the first inequality holds because by assumption 10 that $G(\bar{x}_0)$ contains $G(\underline{x}_0)$ and so contains $\underline{x}_1$. The second inequality holds because by assumption 8 that $u(x_0, x_1)$ is non-decreasing with respect to the components of $x_0$. Hence Q(f) is non-decreasing.

Hence if f is any member of C(X), then $Q^n(f)(x_0) = Q(Q^{n-1}(f))(x_0)$ is a non-decreasing function of $x_0$, for n = 1, 2, .... . Since $V = \lim_{n \to \infty} Q^n(f)$, V is a non-decreasing function of $x_0$ as well. That is, since $V(\underline{x}_0) = \lim_{n \to \infty} Q^n(f)(\underline{x}_0)$ and $V(\overline{x}_0) = \lim_{n \to \infty} Q^n(f)(\overline{x}_0)$, the inequalities

$Q^n(f)(\underline{x}_0) \leq Q^n(f)(\overline{x}_0)$, for all n, imply that $V(\underline{x}_0) \leq V(\overline{x}_0)$. ■

Contraction mappings are used in so many contexts that it is useful to have a criterion that makes it easy to check whether a particular mapping is a contraction. Blackwell's sufficient conditions for a contraction provide such a criterion. The criterion applies to any subset of the set of bounded functions on a particular set X, such as the continuous functions on a compact set X. If X is a set, the set of bounded functions on X is

$$B(X) = \{f: X \to R \mid \text{for some } b > 0, |f(x)| \leq b, \text{ for all } x \in X\}.$$

The sup norm, $\|.\|$, on B(X) is defined by the equation

$$\|f\| = \sup_{x \in X} |f(x)|.$$

<u>Blackwell's Theorem</u> 11.13: Let $T: Z \to Z$, where Z is a subset of B(X), and apply the sup norm $\|.\|$ to points in Z. Suppose that for any $f \in Z$, $f + a \in Z$, for any non-negative number a, where f + a is the function that assigns the number f(x) + a to any $x \in X$. Assume also that

a) if f and g in Z are such that $f(x) \leq g(x)$, for all $x \in X$, then $T(f)(x) \leq T(g)(x)$, for all $x \in X$,

b) there exists $\beta \in (0, 1)$ such that $T(f + a)(x) \leq T(f)(x) + \beta a$, for any $f \in Z$ and any non-negative number a.

Then T is a contraction with modulus $\beta$.

<u>Proof</u>: Let f and g belong to Z. Then $f(x) \leq g(x) + \|f - g\|$, for all $x \in X$, so that

$$T(f)(x) \leq T(g + \|f - g\|)(x) \leq T(g)(x) + \beta\|f - g\|, \tag{11.4}$$

where the first inequality follows from property a and the second from property b. Similarly since $g(x) \leq f(x) + \|f - g\|$, for all $x \in X$

$$T(g)(x) \leq T(f)(x) + \beta\|f - g\|, \tag{11.5}$$

for all $x \in X$. Inequalities 11.4 and 11.5 imply that

$$|T(f)(x) - T(g)(x)| \leq \beta\|f - g\|$$

for all $x \in X$, and hence

$$\|T(f) - T(g)\| \leq \beta\|f - g\|,$$

which means that T is a contraction with modulus $\beta$. ■

It is not hard to check that Blackwell's theorem can be applied to prove that the mapping Q of equation 11.3 is a contraction, which is the content of lemma 11.10

I present a growth model as an economic illustration of the material of this lecture.

Example: (One sector optimal growth model) Consider a world with two commodities, labor and a produced good which can be used for consumption or as an input into production. The output in period t+1 is $y_{t+1} = f(K_t, L_{t+1})$, where f is the production function, $K_t$ is the input of produced good (or capital) in period t, and $L_{t+1}$ is the input of labor in period t+1. The utility enjoyed by society in period t is $v(C_t)$, where v is the utility function and $C_t$ is the consumption of the produced good in period t. The variables $y_t$, $K_t$, $C_t$, and $L_t$ are all non-negative. Assume that the economy is endowed with one unit of labor in every period. A program consists of $(y_t, K_t, C_t, L_t)_{t=0}^{\infty}$, where $y_0$ is given and is the initial stock of output. A program is feasible (with initial stock $y_0$) if

$$L_t \leq 1,$$

$$C_0 + K_0 \leq y_0,$$

and

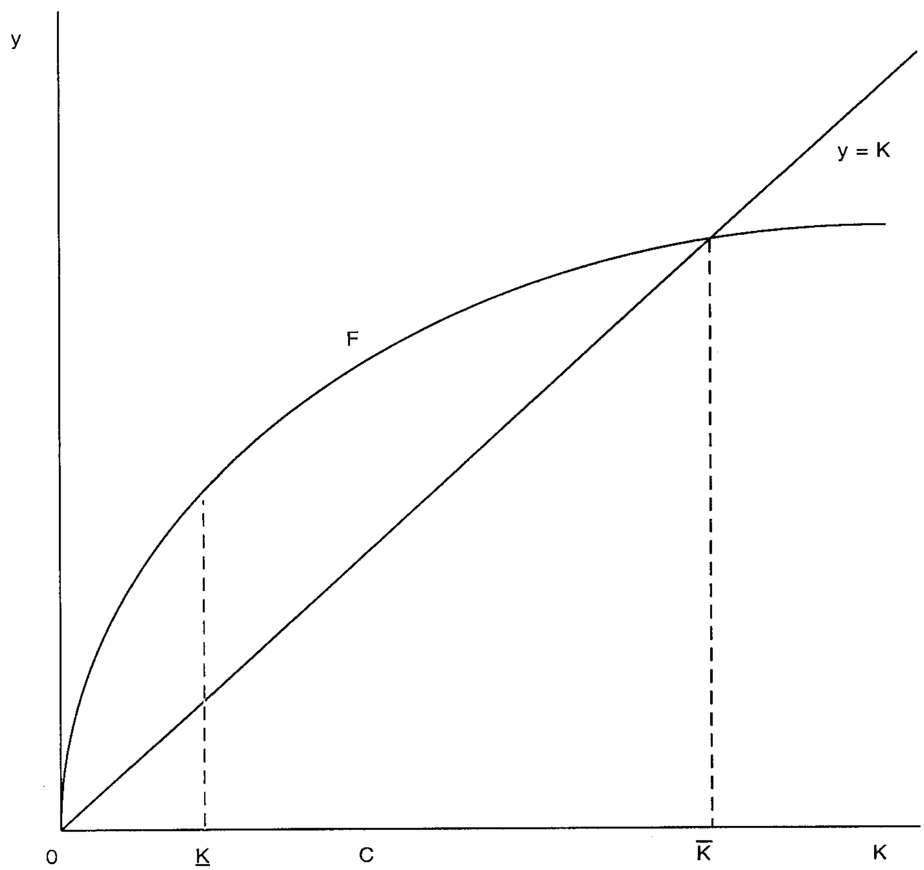
$$C_t + K_t \leq y_t = f(K_{t-1}, L_t),$$

for t = 1, 2, .... . The problem studied is

$$\max_{\substack{(y_t, K_t, C_t, L_t)_{t=0}^{\infty} \\ \text{is feasible with initial} \\ \text{stock } y_0}} \sum_{t=0}^{\infty} \beta^t v(C_t),$$

where $0 < \beta < 1$ and $y_0$ is given and positive. Assume that v is increasing and f(K, L) is non-decreasing and is increasing if K > 0 and L > 0. Then if $(y_t, K_t, C_t, L_t)_{t=0}^{\infty}$ is optimal in the above problem, we may assume that $L_t = 1$, for all t. If we make this assumption, then we may replace the f(K, L) by the function F(K) = f(K, 1), which is increasing, and eliminate the variable L.

In order to relate this example to what we have been doing, let the state at time t be output, $y_t$, and let $u(y_t, y_{t+1}) = v(y_t - F^{-1}(y_{t+1}))$, where $F^{-1}$ is the inverse of F. This inverse exists, because F is increasing. Let $G(y_t) = [0, F(y_t)]$. We need conditions under which there exists a compact set X such that $G(y) \subset X$, if $y \in X$. Assume that f is continuous and homogeneous of degree 1 and that $f(K, 0) = 0 = f(0, L)$, for all K and L. Assume also the for some $\underline{K} > 0$, $f(\underline{K}, 1) > \underline{K}$. Because f is homogeneous of degree 1, $f(K, 1) = Kf(1, K^{-1})$. Because f is continuous and $f(1, 0) = 0$, it follows that $\lim_{K \to \infty} f(1, K^{-1}) = f(1, 0) = 0$. Hence $F(K, 1) = Kf(1, K^{-1}) < K$, if K is large enough. Since $f(K, 1)$ is continuous and $f(\underline{K}, 1) > \underline{K}$, corollary 4.19 of the intermediate value theorem implies that there is a $\overline{K}$ such that $\overline{K} > \underline{K}$ and $f(\overline{K}, 1) = \overline{K}$. Then $f(K, 1) > K$, if $0 < K < \overline{K}$ and $f(K, 1) < K$, if $K > \overline{K}$. If $X = [0, \overline{K}]$, then $G(y) \subset X$, if $y \in X$, as may be seen from the figure below. The set on which u is defined is $A = \{(x, y) \in X \times X \mid y \leq F(x)\}$. This finishes the discussion of the example.



I now interpret theorem 11.3 in the context of the one sector growth model presented in the appendix to lecture 10. Recall that in that model, the state at time t is output, $y_t$, and the utility at that time is $u(y_t, y_{t+1}) = v(y_t - F^{-1}(y_{t+1}))$. Therefore the Euler equation,

$$\frac{\partial}{\partial y_{t+1}} u(y_t, y_{t+1}) = -\beta \frac{\partial}{\partial y_{t+1}} u(y_{t+1}, y_{t+2}),$$

becomes

$$-\frac{d}{dC} v(y_t - F^{-1}(y_{t+1})) \frac{dF^{-1}(y_{t+1})}{dy} = -\beta \frac{d}{dC} v(y_{t+1} - F^{-1}(y_{t+2})). \tag{11.6}$$

If we let $K_t = F^{-1}(y_{t+1})$, $C_t = y_t - K_t = y_t - F^{-1}(y_{t+1})$, and $C_{t+1} = y_{t+1} - F^{-1}(y_{t+2})$, then

$$\frac{dF^{-1}(y_{t+1})}{dy} = \frac{1}{dF(K_t)/dK},$$

and equation 11.6 becomes

$$\frac{dv(C_t)}{dC} = \beta \frac{dv(C_{t+1})}{dC} \frac{dF(K_t)}{dK}. \tag{11.7}$$

This version of the Euler equation makes sense intuitively, for suppose one changes investment in period t by $\Delta K_t$. Then consumption in period t changes by $-\Delta K_t$ and output in period t+1 changes by approximately $\frac{dF(K_t)}{dK}\Delta K_t$. If this change in output is entirely absorbed by change in current consumption, then consumption in period t+1 changes by approximately $\frac{dF(K_t)}{dK}\Delta K_t$. The change in total utility is approximately

$$-\beta^t \frac{dv(C_t)}{dC}\Delta K_t + \beta^{t+1} \frac{dv(C_{t+1})}{dC} \frac{dF(K_t)}{dK}\Delta K_t = \beta^t\left(-\frac{dv(C_t)}{dC} + \beta \frac{dv(C_{t+1})}{dC} \frac{dF(K_t)}{dK}\right)\Delta K_t$$

At an optimal program, this change in total utility must be zero, so that equation 11.7 applies.

Equation 11.7 and the feasibility equation

$$C_{t+1} + K_{t+1} = f(K_t)$$

together determine the evolution of $C_t$ and $K_t$ and hence of $y_t$, for suppose $y_0$ is given. Choose a value for $K_0$ such that $0 \le K_0 \le y_0$. Then $C_0 = y_0 - K_0$. By equation 11.7,

$$\frac{dv(C_1)}{dC} = \left(\beta \frac{dF(K_0)}{dK}\right)^{-1} \frac{dv(C_0)}{dC}.$$

If dv/dC is decreasing, this equation determines $C_1$. Since $K_1 = F(K_0) - C_1$, $K_1$ is determined. We can continue this process by induction on t either indefinitely or until we arrive at a value of t+1 such that $C_{t+1} > F(K_t)$, so that continuation becomes impossible.

I have just shown how in the one sector growth model the Euler equation and feasibility generate through an inductive process a whole program from a single initial condition. I now present a specific example to show that even if this process can be continued indefinitely, the resulting program may not be optimal and is optimal only if it satisfies a transversality condition.