Homework #1
Mathematical Methods II
Spring 2009

1. Dynamic programming involves iterating on the Bellman operator

$$(Tw)(x) = \max_{a \in \Gamma(x)} \{ r(x,a) + \beta w(t(x,a)) \}$$

to convergence. Consider the alternative operator

$$(T_\pi w)(x) = r(x, \pi(x)) + \beta w(t(x, \pi(x))),$$

for some specified policy function $\pi(x)$. Assume that $\beta < 1$, r is bounded and continuous, and t is continuous.

   (a) Prove that T_π is a strict contraction mapping on the space of continuous and bounded functions for any π.

   Since r is bounded we are free to choose $\phi(x) = 1$. Let $v(x) \geq w(x)$ $\forall x$. Then

$$\begin{aligned} (T_\pi v)(x) &= r(x, \pi(x)) + \beta v(t(x, \pi(x))) \\ &\geq r(x, \pi(x)) + \beta w(t(x, \pi(x))) \\ &= (T_\pi w)(x). \end{aligned}$$

   Now let A be a constant function. Then

$$\begin{aligned} (T_\pi(w + A))(x) &= r(x, \pi(x)) + \beta(w(t(x, \pi(x))) + A) \\ &= r(x, \pi(x)) + \beta w(t(x, \pi(x))) + \beta A \\ &= (T_\pi w)(x) + \beta A. \end{aligned}$$

   Having verified both Blackwell-Boyd conditions we prove that T_π is a strict contraction mapping.

   (b) In practice we cannot carry out the Bellman iterations exactly – we must carry out iterations on an approximation to the current w function. Consider approximations L that map functions w into a family of functions F that satisfy two conditions. First, they are **projections**: $L \circ L = L$ (that is, if $w \in F$ then $Lw = w$). Second, they are **nonexpansive**:

$$||Lv - Lw||_\infty \leq ||v - w||_\infty$$

   $\forall v, w$. Define

$$\begin{aligned} \left(\widehat{T}w\right)(x) &= [(L \circ T)(w)](x) \\ &= \left[L\left(\max_{a \in \Gamma(x)} \{ r(x,a) + \beta w(t(x,a)) \} \right) \right](x) \end{aligned}$$

to be the Bellman operator on approximated functions. Let v^* denote the true value function and v_π denote the fixed point of T_π. Show that

$$v^*(x) - v_\pi(x) \leq \frac{2}{(1-\beta)^2}\left(\beta\epsilon + \|v^* - Lv^*\|_\infty\right)$$

for some $\epsilon > 0$.
First, by the triangle inequality we have

$$\|v^* - v_\pi\|_\infty \leq \left\|v^* - \widehat{T}^n v_0\right\|_\infty + \left\|\widehat{T}^n v_0 - v_\pi\right\|_\infty .$$

By the triangle inequality and the contraction mapping theorem, we have

$$\begin{aligned} \|v^* - v_N\|_\infty &\leq \|v^* - Tv_N\|_\infty + \|Tv_N - v_N\|_\infty \\ &\leq \beta\|v^* - v_N\|_\infty + \|Tv_N - v_N\|_\infty \end{aligned}$$

so that

$$(1-\beta)\|v^* - v_N\|_\infty \leq \|Tv_N - v_N\|_\infty .$$

Now

$$\begin{aligned} \|Tv_N - v_N\|_\infty &\leq \|Tv_N - v_{N+1}\|_\infty + \|v_{N+1} - v_N\|_\infty \\ &\leq \|v_{N+1} - Tv_N\|_\infty + \beta\|v_N - v_{N-1}\|_\infty \\ &= \|LTv_N - Tv_N\|_\infty + \beta\epsilon. \end{aligned}$$

Thus, we have

$$(1-\beta)\|v^* - v_N\|_\infty \leq \beta\epsilon + \|LTv_N - Tv_N\|_\infty .$$

Next, by the triangle inequality we have

$$\|v_N - v_\pi\|_\infty \leq \|v_N - Tv_N\|_\infty + \|Tv_N - v_\pi\|_\infty .$$

By the definition of T, we have

$$Tv_N(x) = \max_{a\in\Gamma(x)}\left\{r(x,a) + \beta v_N(t(x,a))\right\}$$

and from the definition of T_π we have

$$T_\pi v_N(x) = r(x,\pi(x)) + \beta v_N(t(x,\pi(x))) .$$

Since π maximizes given v_N, $Tv_N = T_\pi v_N$. Thus, since T_π is a contraction we have

$$\|Tv_N - v_\pi\|_\infty \leq \|T_\pi v_N - T_\pi v_\pi\|_\infty \leq \beta\|v_N - v_\pi\|_\infty .$$

Therefore,

$$||v_N - v_\pi||_\infty \leq ||v_N - Tv_N||_\infty + \beta\,||v_N - v_\pi||_\infty$$

or

$$(1-\beta)\,||v_N - v_\pi||_\infty \leq ||v_N - Tv_N||_\infty\,.$$

By the previous result we can conclude

$$(1-\beta)\,||v_N - v_\pi||_\infty \leq \beta\epsilon + ||LTv_N - Tv_N||_\infty\,.$$

Combining our inequalities leads to

$$(1-\beta)\,||v^* - v_\pi||_\infty \leq 2\left(\beta\epsilon + ||LTv_N - Tv_N||_\infty\right)$$

which can be rearranged to obtain the error bound condition.

(c) Interpret the error bound condition.

The condition states that the difference between the true value function and the value function obtained using the approximation is bounded by a multiple of the distance between the true value function and the approximation to the true value function, letting $\epsilon \to 0$. That multiple is smaller when the operator T is "more contractive" (that is, when β is smaller).

2. Consider the dynamic program

$$v(x) = \max_a \left\{Qx^2 + Ra^2 + 2Wax + \beta v(x')\right\}$$

with law of motion

$$x' = Ax + Ba.$$

Assume that (Q, R, W) are such that the objective function is strictly increasing and strictly concave over the relevant ranges for (x, a). Both x and a are scalars.

(a) Prove that the value function takes the form

$$v(x) = Px^2$$

and the decision rule takes the form

$$a(x) = Fx.$$

Substituting those guesses into the Bellman equation yields

$$Px^2 = \max_a \left\{Qx^2 + Ra^2 + 2Wax + \beta P\left(Ax + Ba\right)^2\right\}.$$

The first-order condition is

$$2Ra + 2Wx + 2B\beta P\left(Ax + Ba\right) = 0.$$

Solving for a yields

$$a\left(x\right) = -\frac{W + \beta BPA}{R + \beta PB^2}x.$$

Inserting this into the Bellman equation yields

$$Px^2 = Qx^2 + R\left(\frac{W + \beta BPA}{R + \beta PB^2}\right)^2 x^2 - 2W\left(\frac{W + \beta BPA}{R + \beta PB^2}\right)x^2 + \beta P\left(A - B\left(\frac{W + \beta BPA}{R + \beta PB^2}\right)\right)^2 x^2.$$

P is therefore a solution to

$$P = Q + R\left(\frac{W + \beta BPA}{R + \beta PB^2}\right)^2 - 2W\left(\frac{W + \beta BPA}{R + \beta PB^2}\right) + \beta P\left(A - B\left(\frac{W + \beta BPA}{R + \beta PB^2}\right)\right)^2.$$

(b) Does P exist? Is P unique? What is the sign of P?

To prove that P exists and must be unique, we only need to show that iterations on the Bellman operator produce a contraction on a Banach space. It is clear that if $v\left(x\right)$ is quadratic then $\left(Tv\right)\left(x\right)$ is also quadratic and that the limit of a sequence of quadratic functions is also quadratic; since quadratic functions are ϕ-bounded by the constant function, we only need to prove that the operation is monotone and contracts constant functions. Monotonicity: let P_1, P_2 be such that $P_1x^2 \geq P_2x^2\ \forall x$. Then

$$\begin{aligned}\left(TP_1\right)\left(x\right) &= \max\left\{Qx^2 + Ra^2 + 2Wax + \beta P_1\left(x'\right)^2\right\}\\ &\geq \max\left\{Qx^2 + Ra^2 + 2Wax + \beta P_2\left(x'\right)^2\right\}\\ &= \left(TP_2\right)\left(x\right).\end{aligned}$$

Discounting constant functions: let Z be a constant function. Then

$$\begin{aligned}\left(T\left(P + Z\right)\right)\left(x\right) &= \max\left\{Qx^2 + Ra^2 + 2Wax + \beta\left(P\left(x'\right)^2 + Z\right)\right\}\\ &= \max\left\{Qx^2 + Ra^2 + 2Wax + \beta P\left(x'\right)^2\right\} + \beta Z\\ &= \left(TP\right)\left(x\right) + \beta Z.\end{aligned}$$

Thus, the Bellman operator is a contraction, so P exists and must be unique. The Bellman operator preserves concavity, so P must be negative since we can start the iterations with $P = 0$ (a concave function).

(c) Assume that $Q = -1$, $R = -2$, and $W = 0$. Let $\beta = 0.96$, $A = 0.5$, and $B = 0.1$. Compute P and F, where

$$a^*\left(x\right) = -Fx.$$

Iterating on the Bellman operator yields the solution $P = -1.3132$ and

$$\begin{aligned} F &= \frac{W + \beta B P A}{R + \beta P B^2} \\ &= 0.0313. \end{aligned}$$