Lecture Notes 10: Dynamic Programming
Part A: Stochastic Difference Equations

Peter J. Hammond

2016 September 30th
Lecture Outline

Linear Equations in One Variable

Optimal Saving

The Two Period Problem

The $T$ Period Problem

A General Problem

Infinite Time Horizon

Main Theorem

Policy Improvement

Unboundedness
A simple stochastic linear difference equation of the first order in one variable takes the form

\[ x_t = ax_{t-1} + \epsilon_t \quad (t \in \mathbb{N}) \]

Here \( a \) is a real parameter, and each \( \epsilon_t \) is a real random disturbance.

Assume that:

1. there is a given or pre-determined initial state \( x_0 \);
2. the random variables \( \epsilon_t \)
   are independent and identically distributed (IID)
   with mean \( \mathbb{E}\epsilon_t = 0 \) and variance \( \mathbb{E}\epsilon_t^2 = \sigma^2 \).

A special case is when the disturbances are all normally distributed — i.e., \( \epsilon_t \sim N(0, \sigma^2) \).
Explicit Solution and Conditional Mean

For each fixed outcome \( \epsilon^N = (\epsilon_t)_{t \in \mathbb{N}} \) of the random sequence, there is a unique solution which can be written as

\[
x_t = a^t x_0 + \sum_{s=1}^{t} a^{t-s} \epsilon_s
\]

The main stable case occurs when \( |a| < 1 \).

Then each term of the sum \( a^t x_0 + \sum_{s=1}^{t} a^{t-s} \epsilon_s \) converges to 0 as \( t \to \infty \).

This is what econometricians or statisticians call a first-order autoregressive (or AR(1)) process.

In fact, given \( x_0 \) at time 0, our assumption that \( \mathbb{E}[\epsilon_s] = 0 \) for all \( s = 1, 2, \ldots, t \) implies that the conditional mean of \( x_t \) is

\[
m_t := \mathbb{E}[x_t | x_0] = \mathbb{E} \left[ a^t x_0 + \sum_{s=1}^{t} a^{t-s} \epsilon_s | x_0 \right] = a^t x_0
\]
Conditional Variance

The conditional variance, however, is given by

\[ v_t := \mathbb{E} \left[ (x_t - m_t)^2 \mid x_0 \right] = \mathbb{E}[(x_t - a^t x_0)^2 \mid x_0] = \mathbb{E} \left[ \sum_{s=1}^{t} a^{t-s} \epsilon_s \right]^2 \]

In the case we are considering with independently distributed disturbances \( \epsilon_s \), the variance of a sum is the sum of the variances.

Hence

\[ v_t = \sum_{s=1}^{t} \mathbb{E} \left[ a^{t-s} \epsilon_s \right]^2 = \sum_{s=1}^{t} a^{2(t-s)} \mathbb{E} \epsilon_s^2 = \sigma^2 \sum_{s=1}^{t} a^{2(t-s)} \]

Using the rule for summing the geometric series \( \sum_{s=1}^{t} a^{2(t-s)} \), we finally obtain

\[ v_t = \frac{1 - a^{2t}}{1 - a^2} \sigma^2 \]
Recall that if $X \sim N(\mu, \sigma^2)$, then the characteristic function defined by $\phi_X(t) = \mathbb{E}[e^{iXt}]$ takes the form

$$\phi_X(t) = \mathbb{E}[e^{iXt}] = \int_{-\infty}^{+\infty} e^{ixt} \frac{1}{\sqrt{2\pi\sigma^2}} \exp \left( -\frac{1}{2} \left( \frac{x - \mu}{\sigma} \right)^2 \right) dx$$

This reduces to $\phi_X(t) = \exp \left( it\mu - \frac{1}{2} \sigma^2 t^2 \right)$.

Hence, if $Z = X + Y$ where $X \sim N(\mu_X, \sigma_X^2)$ and $Y \sim N(\mu_Y, \sigma_Y^2)$ are independent random variables, then

$$\phi_Z(t) = \mathbb{E}[e^{iZt}] = \mathbb{E}[e^{i(X+Y)t}] = \mathbb{E}[e^{iXt}e^{iYt}] = \mathbb{E}[e^{iXt}]\mathbb{E}[e^{iYt}]$$
So
\[ \phi_Z(t) = \exp \left( it\mu_X - \frac{1}{2}\sigma_X^2 t^2 \right) \exp \left( it\mu_Y - \frac{1}{2}\sigma_Y^2 t^2 \right) \]
\[ = \exp \left( it(\mu_X + \mu_Y) - \frac{1}{2}(\sigma_X^2 + \sigma_Y^2) t^2 \right) \]
\[ = \exp \left( it\mu_Z - \frac{1}{2}\sigma_Z^2 t^2 \right) \]

where \( \mu_Z = \mu_X + \mu_Y = E(X + Y) \) is the mean of \( X + Y \),
and \( \sigma_Z^2 = \sigma_X^2 + \sigma_Y^2 \) is the variance of \( X + Y \).

It follows that \( t \mapsto \phi_Z(t) \)
is the characteristic function of a random variable \( Z \sim N(\mu_Z, \sigma_Z^2) \)
where \( \mu_Z = \mu_X + \mu_Y = E(X + Y) \) and \( \sigma_Z^2 = \sigma_X^2 + \sigma_Y^2 \).

That is, the sum \( Z = X + Y \)
of two independent normally distributed random variables \( X \) and \( Y \)
is also normally distributed, with:

1. mean equal to the sum of the means;
2. variance equal to the sum of the variances.
The Gaussian Case and the Asymptotic Distribution

In the particular case when each $\epsilon_t$ is normally distributed as well as IID, then $x_t$ is also normally distributed with mean $m_t$ and variance $\nu_t$.

As $t \to \infty$, the conditional mean $m_t = a^t x_0 \to 0$ and the conditional variance

$$
\nu_t = \frac{1 - a^{2t}}{1 - a^2} \sigma^2 \to \nu := \frac{\sigma^2}{1 - a^2}
$$

In the case when each $\epsilon_t$ is normally distributed, this implies that the asymptotic distribution of $x_t$ is also normal, with mean 0 and variance $\nu = \sigma^2/(1 - a^2)$.
Stationarity

Now suppose that $x_0$ itself has this asymptotic normal distribution — suppose that $x_0 \sim N(0, \sigma^2/(1 - a^2))$.

This is what the distribution of $x_0$ would be if the process had started at $t = -\infty$ instead of at $t = 0$.

Then the unconditional mean of each $x_t$ is $E x_t = a^t E x_0 = 0$.

On the other hand, because $x_{t+k} = a^k x_t + \sum_{s=1}^{k} a^{k-s} \epsilon_{t+s}$, the unconditional covariance of $x_t$ and $x_{t+k}$ is

$$E(x_{t+k}x_t) = E[a^k x_t^2] = a^k \nu = \frac{a^k}{1 - a^2} \sigma^2 \quad (k = 0, 1, 2 \ldots)$$

In fact, given any $t$, the joint distribution of the $r$ random variables $x_t, x_{t+1}, \ldots, x_{t+r-1}$ is multivariate normal with variance–covariance matrix having elements $E(x_{t+k}x_t) = a^k \sigma^2/(1 - a^2)$, independent of $t$.

Because of this independence, the process is said to be stationary.
Lecture Outline

Linear Equations in One Variable

Optimal Saving

The Two Period Problem

The $T$ Period Problem

A General Problem

Infinite Time Horizon

Main Theorem

Policy Improvement

Unboundedness
Intertemporal Utility

Consider a household which at time $s$ is planning its intertemporal consumption stream $c^T_s := (c_s, c_{s+1}, \ldots, c_T)$ over periods $t$ in the set $\{s, s+1, \ldots, T\}$.

Its intertemporal utility function $\mathbb{R}^{T-s+1} \ni c^T_s \mapsto U^T_s(c^T_s) \in \mathbb{R}$ is assumed to take the additively separable form

$$U^T_s(c^T_s) := \sum_{t=s}^T u_t(c_t)$$

where the one period felicity functions $c \mapsto u_t(c)$ are differentiably increasing and strictly concave (DISC) — i.e., $u'_t(c) > 0$, and $u''_t(c) < 0$ for all $t$ and all $c > 0$.

As before, the household faces:

1. fixed initial wealth $w_s$;
2. a terminal wealth constraint $w_{T+1} \geq 0$. 
Also as before, we assume a wealth accumulation equation \( w_{t+1} = \tilde{r}_t (w_t - c_t) \), where \( \tilde{r}_t \) is the household’s gross rate of return on its wealth in period \( t \).

It is assumed that:

1. the return \( \tilde{r}_t \) in each period \( t \) is a random variable with positive values;

2. the return distributions for different times \( t \) are stochastically independent;

3. starting with predetermined wealth \( w_s \) at time \( s \), the household seeks to maximize the expectation \( \mathbb{E}_s [ U_s^T (c_s^T) ] \) of its intertemporal utility.
Lecture Outline

Linear Equations in One Variable

Optimal Saving

The Two Period Problem

The $T$ Period Problem

A General Problem

Infinite Time Horizon

Main Theorem

Policy Improvement

Unboundedness
Two Period Case

We work backwards from the last period, when \( s = T \).

In this last period the household will obviously choose \( c_T = w_T \), yielding a maximized utility equal to \( V_T(w_T) = u_T(w_T) \).

Next, consider the penultimate period, when \( s = T - 1 \). The consumer will want to choose \( c_{T-1} \) in order to maximize

\[
\begin{align*}
    & u_{T-1}(c_{T-1}) + \mathbb{E}_{T-1} V_T(w_T) \\
    & \text{subject to the wealth constraint}
\end{align*}
\]

\[
\begin{align*}
    w_T = \tilde{r}_{T-1} (w_{T-1} - c_{T-1}) \\
    \text{random gross return saving}
\end{align*}
\]

result of an optimal policy in period \( T \)
First-Order Condition

Substituting both the function $V_T(w_T) = u_T(w_T)$ and the wealth constraint into the objective reduces the problem to

$$\max_{c_{T-1}} \{ u_{T-1}(c_{T-1}) + \mathbb{E}_{T-1}[u_T(\tilde{r}_{T-1}(w_{T-1} - c_{T-1}))] \}$$

subject to $0 \leq c_{T-1} \leq w_{T-1}$ and $\tilde{c}_T := \tilde{r}_{T-1}(w_{T-1} - c_{T-1})$.

Assume we can differentiate under the integral sign, and that there is an interior solution with $0 < c_{T-1} < w_{T-1}$.

Then the first-order condition (FOC) is

$$0 = u'_{T-1}(c_{T-1}) + \mathbb{E}_{T-1}[(−\tilde{r}_{T-1})u'_{T}(\tilde{r}_{T-1}(w_{T-1} - c_{T-1}))]$$
The Stochastic Euler Equation

Rearranging the first-order condition while recognizing that $\tilde{c}_T := \tilde{r}_{T-1}(w_{T-1} - c_{T-1})$, one obtains

$$u'_{T-1}(c_{T-1}) = \mathbb{E}_{T-1}[\tilde{r}_{T-1} u'_{T}(\tilde{r}_{T-1}(w_{T-1} - c_{T-1}))]$$

Dividing by $u'_{T-1}(c_{T-1})$ gives the stochastic Euler equation

$$1 = \mathbb{E}_{T-1} \left[ \tilde{r}_{T-1} \frac{u'_T(\tilde{c}_T)}{u'_{T-1}(c_{T-1})} \right] = \mathbb{E}_{T-1} \left[ \tilde{r}_{T-1} \text{MRS}^T_{T-1}(c_{T-1}; \tilde{c}_T) \right]$$

involving the marginal rate of substitution function

$$\text{MRS}^T_{T-1}(c_{T-1}; \tilde{c}_T) := \frac{u'_T(\tilde{c}_T)}{u'_{T-1}(c_{T-1})}$$
The CES Case

For the marginal utility function \( c \mapsto u'(c) \), its \textit{elasticity of substitution} is defined for all \( c > 0 \) by \( \eta(c) := d \ln u'(c) / d \ln c \).

Then \( \eta(c) \) is both the degree of relative risk aversion, and the degree of relative fluctuation aversion.

A \textit{constant elasticity of substitution} (or CES) utility function satisfies \( d \ln u'(c) / d \ln c = -\epsilon < 0 \) for all \( c > 0 \).

The marginal rate of substitution satisfies \( u'(c) / u'(\bar{c}) = (c/\bar{c})^{-\epsilon} \) for all \( c, \bar{c} > 0 \).
Normalized Utility

Normalize by putting $u'(1) = 1$, implying that $u'(c) \equiv c^{-\epsilon}$.

Then integrating gives

$$u(c; \epsilon) = u(1) + \int_1^c x^{-\epsilon} dx$$

$$= \begin{cases} 
  u(1) + \frac{c^{1-\epsilon} - 1}{1 - \epsilon} & \text{if } \epsilon \neq 1 \\
  u(1) + \ln c & \text{if } \epsilon = 1 
\end{cases}$$

Introduce the final normalization

$$u(1) = \begin{cases} 
  1 & \text{if } \epsilon \neq 1 \\
  0 & \text{if } \epsilon = 1 
\end{cases}$$

The utility function is reduced to

$$u(c; \epsilon) = \begin{cases} 
  \frac{c^{1-\epsilon} - 1}{1 - \epsilon} & \text{if } \epsilon \neq 1 \\
  \ln c & \text{if } \epsilon = 1 
\end{cases}$$
The Stochastic Euler Equation in the CES Case

Consider the CES case when $u'_t(c) \equiv \delta_t c^{-\epsilon}$, where each $\delta_t$ is the discount factor for period $t$.

Definition

The one-period discount factor in period $t$ is defined as $\beta_t := \delta_{t+1}/\delta_t$.

Then the stochastic Euler equation takes the form

$$1 = \mathbb{E}_{T-1} \left[ \tilde{r}_{T-1} \beta_{T-1} \left( \frac{\tilde{c}_T}{c_{T-1}} \right)^{-\epsilon} \right]$$

Because $c_{T-1}$ is being chosen at time $T - 1$, this implies that

$$(c_{T-1})^{-\epsilon} = \mathbb{E}_{T-1} \left[ \tilde{r}_{T-1} \beta_{T-1} (\tilde{c}_T)^{-\epsilon} \right]$$
The Two Period Problem in the CES Case

In the two-period case, we know that

$$\tilde{c}_T = \tilde{w}_T = \tilde{r}_{T-1}(w_{T-1} - c_{T-1})$$

in the last period, so the Euler equation becomes

$$(c_{T-1})^{-\epsilon} = E_{T-1} \left[ \tilde{r}_{T-1} \beta_{T-1} (\tilde{c}_T)^{-\epsilon} \right]$$

$$= \beta_{T-1}(w_{T-1} - c_{T-1})^{-\epsilon} E_{T-1} \left[ (\tilde{r}_{T-1})^{1-\epsilon} \right]$$

Take the $(-1/\epsilon)$ th power of each side and define

$$\rho_{T-1} := (\beta_{T-1} E_{T-1} \left[ (\tilde{r}_{T-1})^{1-\epsilon} \right])^{-1/\epsilon}$$

to reduce the Euler equation to $c_{T-1} = \rho_{T-1}(w_{T-1} - c_{T-1})$ whose solution is evidently $c_{T-1} = \gamma_{T-1} w_{T-1}$ where

$$\gamma_{T-1} := \rho_{T-1}/(1 + \rho_{T-1}) \quad \text{and} \quad 1 - \gamma_{T-1} = 1/(1 + \rho_{T-1})$$

are respectively the optimal consumption and savings ratios.

It follows that $\rho_{T-1} = \gamma_{T-1}/(1 - \gamma_{T-1})$ is the consumption/savings ratio.
The optimal policy in periods $T$ and $T - 1$ is $c_t = \gamma_t w_t$ where $\gamma_T = 1$ and $\gamma_{T-1}$ has just been defined.

In this CES case, the discounted utility of consumption in period $T$ is $V_T(w_T) := \delta_T u(w_T; \epsilon)$.

The discounted expected utility at time $T - 1$ of consumption in periods $T$ and $T - 1$ together is

$$V_{T-1}(w_{T-1}) = \delta_{T-1} u(\gamma_{T-1} w_{T-1}; \epsilon) + \delta_T \mathbb{E}_{T-1}[u(\tilde{w}_T; \epsilon)]$$

where $\tilde{w}_T = \tilde{r}_{T-1}(1 - \gamma_{T-1}) w_{T-1}$.
Discounted Expected Utility in the Logarithmic Case

In the logarithmic case when $\epsilon = 1$, one has

$$V_{T-1}(w_{T-1}) = \delta_{T-1} \ln(\gamma_{T-1} w_{T-1}) + \delta_T \mathbb{E}_{T-1} [\ln (\tilde{r}_{T-1} (1 - \gamma_{T-1}) w_{T-1})]$$

It follows that

$$V_{T-1}(w_{T-1}) = \alpha_{T-1} + (\delta_{T-1} + \delta_T) u(w_{T-1}; \epsilon)$$

where

$$\alpha_{T-1} := \delta_{T-1} \ln \gamma_{T-1} + \delta_T \{ \ln(1 - \gamma_{T-1}) + \mathbb{E}_{T-1} [\ln \tilde{r}_{T-1}] \}$$
Discounted Expected Utility in the CES Case

In the CES case when $\epsilon \neq 1$, one has

$$(1 - \epsilon) V_{T-1}(w_{T-1}) = \delta_{T-1}(\gamma_{T-1} w_{T-1})^{1-\epsilon}$$

$$+ \delta_T [(1 - \gamma_{T-1}) w_{T-1}]^{1-\epsilon} \mathbb{E}_{T-1}[(\tilde{r}_{T-1})^{1-\epsilon}]$$

so $V_{T-1}(w_{T-1}) = v_{T-1} u(w_{T-1}; \epsilon)$ where

$$v_{T-1} := \delta_{T-1}(\gamma_{T-1})^{1-\epsilon} + \delta_T (1 - \gamma_{T-1})^{1-\epsilon} \mathbb{E}_{T-1}[(\tilde{r}_{T-1})^{1-\epsilon}]$$

In both cases, one can write $V_{T-1}(w_{T-1}) = \alpha_{T-1} + v_{T-1} u(w_{T-1}; \epsilon)$ for a suitable additive constant $\alpha_{T-1}$ (which is 0 in the CES case) and a suitable multiplicative constant $v_{T-1}$.
Lecture Outline

Linear Equations in One Variable

Optimal Saving

The Two Period Problem

The $T$ Period Problem

A General Problem

Infinite Time Horizon

Main Theorem

Policy Improvement

Unboundedness
The Time Line

In each period $t$, suppose:

- the consumer starts with known wealth $w_t$;
- then the consumer chooses consumption $c_t$, along with savings or residual wealth $w_t - c_t$;
- there is a cumulative distribution function $F_t(r)$ on $\mathbb{R}$ that determines the gross return $\tilde{r}_t$ as a positive-valued random variable.

After these three steps have been completed, the problem starts again in period $t + 1$, with the consumer’s wealth known to be $w_{t+1} = \tilde{r}_t(w_t - c_t)$. 
Expected Conditionally Expected Utility

Starting at any $t$, suppose the consumer’s choices, together with the random returns, jointly determine a cdf $F_T^T$ over the space of intertemporal consumption streams $c_T^T$.

The associated expected utility is $E_t \left[ U_T^T (c_T^T) \right]$, using the shorthand $E_t$ to denote integration w.r.t. the cdf $F_T^T$.

Then, given that the consumer has chosen $c_t$ at time $t$, let $E_{t+1}[\cdot|c_t]$ denote the conditional expected utility.

This is found by integrating w.r.t. the conditional cdf $F_{t+1}^T(c_{t+1}^T|c_t)$.

The law of iterated expectations allows us to write the unconditional expectation $E_t \left[ U_T^T (c_T^T) \right]$ as the expectation $E_t[E_{t+1}[U_T^T(c_T^T)|c_t]]$ of the conditional expectation.
The Expectation of Additively Separable Utility

Our hypothesis is that the intertemporal von Neumann–Morgenstern utility function takes the additively separable form

\[ U_t^T (c_t^T) = \sum_{\tau=t}^{T} u_\tau (c_\tau) \]

The conditional expectation given \( c_t \) must then be

\[ \mathbb{E}_{t+1} [U_t^T (c_t^T)|c_t] = u_t (c_t) + \mathbb{E}_{t+1} \left[ \sum_{\tau=t+1}^{T} u_\tau (c_\tau)|c_t \right] \]

whose expectation is

\[ \mathbb{E}_t \left[ \sum_{\tau=t}^{T} u_\tau (c_\tau) \right] = u_t (c_t) + \mathbb{E}_t \left[ \mathbb{E}_{t+1} \left[ \sum_{\tau=t+1}^{T} u_\tau (c_\tau) \right] |c_t \right] \]
The Continuation Value

Let $V_{t+1}(w_{t+1})$ be the state valuation function expressing the maximum of the continuation value

$$
E_{t+1} \left[ U_{t+1}^T(c_{t+1}^T)| w_{t+1} \right] = E_{t+1} \left[ \sum_{\tau=t+1}^{T} u_{\tau}(c_{\tau})| w_{t+1} \right]
$$

as a function of the wealth level or state $w_{t+1}$.

Assume this maximum value is achieved by following an optimal policy from period $t + 1$ on.

Then total expected utility at time $t$ will then reduce to

$$
E_t \left[ U_t^T(\tilde{c}_t^T)| c_t \right] = u_t(c_t) + E_t \left[ E_{t+1} \left[ \sum_{\tau=t+1}^{T} u_{\tau}(c_{\tau})| w_{t+1} \right] | c_t \right] \\
= u_t(c_t) + E_t [ V_{t+1}(\tilde{w}_{t+1})| c_t ] \\
= u_t(c_t) + E_t [ V_{t+1}(\tilde{r}_t(w_t - c_t)) ]
$$
The Principle of Optimality

Maximizing $\mathbb{E}_s \left[ U_s^T (c_s^T) \right]$ w.r.t. $c_s$, taking as fixed the optimal consumption plans $c_t(w_t)$ at times $t = s + 1, \ldots, T$, therefore requires choosing $c_s$ to maximize

$$u_s(c_s) + \mathbb{E}_s[V_{s+1}(\tilde{r}_s(w_s - c_s))]$$

Let $c^*_s(w_s)$ denote a solution to this maximization problem.

Then the value of an optimal plan $(c^*_t(w_t))^T_{t=s}$ that starts with wealth $w_s$ at time $s$ is

$$V_s(w_s) := u_s(c^*_s(w_s)) + \mathbb{E}_s[V_{s+1}(\tilde{r}_s(w_s - c^*_s(w_s)))]$$

Together, these two properties can be expressed as

$$V_s(w_s) = \begin{cases} \max_{0 \leq c_s \leq w_s} \{ u_s(c_s) + \mathbb{E}_s[V_{s+1}(\tilde{r}_s(w_s - c_s))] \} \\ c^*_s(w_s) = \arg \max \{ u_s(c_s) + \mathbb{E}_s[V_{s+1}(\tilde{r}_s(w_s - c_s))] \} \end{cases}$$

which can be described as the **principle of optimality**.
An Induction Hypothesis

Consider once again the case when $u_t(c) \equiv \delta_t u(c; \epsilon)$ for the CES (or logarithmic) utility function that satisfies $u'(c; \epsilon) \equiv c^{-\epsilon}$ and, specifically

$$u(c; \epsilon) = \begin{cases} \frac{c^{1-\epsilon}}{(1 - \epsilon)} & \text{if } \epsilon \neq 1; \\ \ln c & \text{if } \epsilon = 1. \end{cases}$$

Inspired by the solution we have already found for the final period $T$ and penultimate period $T - 1$, we adopt the induction hypothesis that there are constants $\alpha_t, \gamma_t, \nu_t$ ($t = T, T - 1, \ldots, s + 1, s$) for which

$$c_t^*(w_t) = \gamma_t w_t \quad \text{and} \quad V_t(w_t) = \alpha_t + \nu_t u(w_t; \epsilon)$$

In particular, the consumption ratio $\gamma_t$ and savings ratio $1 - \gamma_t$ are both independent of the wealth level $w_t$. 
Applying Backward Induction

Under the induction hypotheses that

\[ c_t^*(w_t) = \gamma_t w_t \quad \text{and} \quad V_t(w_t) = \alpha_t + v_t u(w_t; \epsilon) \]

the maximand

\[ u_s(c_s) + \mathbb{E}_s[V_{s+1}(\tilde{r}_s(w_s - c_s))] \]

takes the form

\[ \delta_s u(c_s; \epsilon) + \mathbb{E}_s[\alpha_{s+1} + v_{s+1} u(\tilde{r}_s(w_s - c_s); \epsilon)] \]

The first-order condition for this to be maximized w.r.t. \( c_s \) is

\[ 0 = \delta_s u'(c_s; \epsilon) - v_{s+1} \mathbb{E}_s[\tilde{r}_s u'(\tilde{r}_s(w_s - c_s); \epsilon)] \]

or, equivalently, that

\[ \tilde{\delta}_s(c_s)^{-\epsilon} = v_{s+1} \mathbb{E}_s[\tilde{r}_s(\tilde{r}_s(w_s - c_s))^{-\epsilon}] = v_{s+1} (w_s - c_s)^{-\epsilon} \mathbb{E}_s[(\tilde{r}_s)^{1-\epsilon}] \]
Solving the Logarithmic Case

When $\epsilon = 1$ and so $u(c; \epsilon) = \ln c$, the first-order condition reduces to $\delta_s(c_s)^{-1} = \nu_{s+1}(w_s - c_s)^{-1}$. Its solution is indeed $c_s = \gamma_s w_s$ where $\delta_s(\gamma_s)^{-1} = \nu_{s+1}(1 - \gamma_s)^{-1}$, implying that $\gamma_s = \delta_s / (\delta_s + \nu_{s+1})$.

The state valuation function then becomes

$$V_s(w_s) = \delta_s u(\gamma_s w_s; \epsilon) + \alpha_{s+1} + \nu_{s+1}\mathbb{E}_s[u(\tilde{r}_s(1 - \gamma_s)w_s; \epsilon)]$$

$$= \delta_s \ln(\gamma_s w_s) + \alpha_{s+1} + \nu_{s+1}\mathbb{E}_s[\ln(\tilde{r}_s(1 - \gamma_s)w_s)]$$

$$= \delta_s \ln(\gamma_s w_s) + \alpha_{s+1} + \nu_{s+1}\{\ln(1 - \gamma_s)w_s + \ln R_s\}$$

where we define the geometric mean certainty equivalent return $R_s$ so that $\ln R_s := \mathbb{E}_s[\ln(\tilde{r}_s)]$. 
The State Valuation Function

The formula

$$V_s(w_s) = \delta_s \ln(\gamma_s w_s) + \alpha_{s+1} + \nu_{s+1}\{\ln(1 - \gamma_s)w_s + \ln R_s\}$$

reduces to the desired form $V_s(w_s) = \alpha_s + \nu_s \ln w_s$ provided we take $\nu_s := \delta_s + \nu_{s+1}$, which implies that $\gamma_s = \delta_s/\nu_s$, and also

$$\alpha_s := \delta_s \ln \gamma_s + \alpha_{s+1} + \nu_{s+1}\{\ln(1 - \gamma_s) + \ln R_s\}$$

$$= \delta_s \ln(\delta_s/\nu_s) + \alpha_{s+1} + \nu_{s+1}\{\ln(\nu_{s+1}/\nu_s) + \ln R_s\}$$

$$= \delta_s \ln \delta_s + \alpha_{s+1} - \nu_s \ln \nu_s + \nu_{s+1}\{\ln \nu_{s+1} + \ln R_s\}$$

This confirms the induction hypothesis for the logarithmic case.

The relevant constants $\nu_s$ are found by summing backwards, starting with $\nu_T = \delta_T$, implying that $\nu_s = \sum_{T=s}^{T} \delta_s$. 
The Stationary Logarithmic Case

In the stationary logarithmic case:

- the felicity function in each period $t$ is $\beta^t \ln c_t$, so the one period discount factor is the constant $\beta$;
- the certainty equivalent return $R_t$ is also a constant $R$.

Then $v_s = \sum_{T=s}^{T} \delta_s = \sum_{T=s}^{T} \beta^\tau = (\beta^s - \beta^{T+1})/(1 - \beta)$, implying that $\gamma_s = \beta^s / v_s = \beta^s (1 - \beta) / (\beta^s - \beta^{T+1})$.

It follows that

$$c_s = \gamma_s w_s = \frac{(1 - \beta) w_s}{1 - \beta^{T-s+1}} = \frac{(1 - \beta) w_s}{1 - \beta^{H+1}}$$

when there are $H := T - s$ periods left before the horizon $T$.

As $H \to \infty$, this solution converges to $c_s = (1 - \beta) w_s$, so the savings ratio equals the constant discount factor $\beta$.

Remarkably, this is also independent of the gross return to saving.
First-Order Condition in the CES Case

Recall that the first-order condition in the CES Case is

$$\delta_s(c_s)^{-\epsilon} = v_{s+1}(w_s - c_s)^{-\epsilon} \mathbb{E}_s[(\tilde{r}_s)^{1-\epsilon}] = v_{s+1}(w_s - c_s)^{-\epsilon} R_s^{1-\epsilon}$$

where we have defined the certainty equivalent return $R_s$ as the solution to $R_s^{1-\epsilon} := \mathbb{E}_s[(\tilde{r}_s)^{1-\epsilon}]$.

The first-order condition indeed implies that $c_s^*(w_s) = \gamma_s w_s$, where $\delta_s(\gamma_s)^{-\epsilon} = v_{s+1}(1 - \gamma_s)^{-\epsilon} R_s^{1-\epsilon}$.

This implies that

$$\gamma_s \frac{1}{1 - \gamma_s} = \left(v_{s+1} R_s^{1-\epsilon} / \delta_s\right)^{-1/\epsilon}$$

or

$$\gamma_s = \frac{\left(v_{s+1} R_s^{1-\epsilon} / \delta_s\right)^{-1/\epsilon}}{1 + \left(v_{s+1} R_s^{1-\epsilon} / \delta_s\right)^{-1/\epsilon}} = \frac{\left(v_{s+1} R_s^{1-\epsilon}\right)^{-1/\epsilon}}{(\delta_s)^{-1/\epsilon} + \left(v_{s+1} R_s^{1-\epsilon}\right)^{-1/\epsilon}}$$
Completing the Solution in the CES Case

Under the induction hypothesis that $V_{s+1}(w) = v_{s+1}w^{1-\epsilon}/(1-\epsilon)$, one also has

$$(1 - \epsilon) V_s(w_s) = \delta_s (\gamma_s w_s)^{1-\epsilon} + v_{s+1} \mathbb{E}_s[(\tilde{r}_s (1 - \gamma_s) w_s)^{1-\epsilon}]$$

This reduces to the desired form $(1 - \epsilon) V_s(w_s) = v_s(w_s)^{1-\epsilon}$, where

$$v_s := \delta_s (\gamma_s)^{1-\epsilon} + v_{s+1} \mathbb{E}_s[(\tilde{r}_s)^{1-\epsilon}](1 - \gamma_s)^{1-\epsilon}$$

$$= \frac{\delta_s (v_{s+1} R_s^{1-\epsilon})^{1-1/\epsilon} + v_{s+1} R_s^{1-\epsilon} (\delta_s)^{1-1/\epsilon}}{[(\delta_s)^{-1/\epsilon} + (v_{s+1} R_s^{1-\epsilon})^{-1/\epsilon}]^{1-\epsilon}}$$

$$= \delta_s v_{s+1} R_s^{1-\epsilon} \frac{(v_{s+1} R_s^{1-\epsilon})^{-1/\epsilon} + (\delta_s)^{-1/\epsilon}}{[(\delta_s)^{-1/\epsilon} + (v_{s+1} R_s^{1-\epsilon})^{-1/\epsilon}]^{1-\epsilon}}$$

$$= \delta_s v_{s+1} R_s^{1-\epsilon} [(\delta_s)^{-1/\epsilon} + (v_{s+1} R_s^{1-\epsilon})^{-1/\epsilon}]^{1-\epsilon}$$

This confirms the induction hypothesis for the CES case.

Again, the relevant constants are found by working backwards.
Histories and Strategies

For each time \( t = s, s + 1, \ldots, T \) between the start \( s \) and the horizon \( T \), let \( h^t \) denote a known history \((w_\tau, c_\tau, \tilde{r}_\tau)^{t}_{\tau=s}\) of the triples \((w_\tau, c_\tau, \tilde{r}_\tau)\) at successive times \( \tau = s, s + 1, \ldots, t \) up to time \( t \).

A general policy the consumer can choose involves a measurable function \( h^t \mapsto \psi_t(h^t) \) mapping each known history up to time \( t \), which determines the consumer’s information set, into a consumption level at that time.

The collection of successive functions \( \psi^T_s = \langle \psi_t \rangle^{T}_{t=s} \) is what a game theorist would call the consumer’s strategy in the extensive form game “against nature”.
Markov Strategies

We found an optimal solution for the two-period problem when $t = T - 1$.

It took the form of a Markov strategy $\psi_t(h^t) := c_t^*(w_t)$, which depends only on $w_t$ as the particular state variable.

The following analysis will demonstrate in particular that at each time $t = s, s + 1, \ldots, T$, under the induction hypothesis that the consumer will follow a Markov strategy in periods $\tau = t + 1, t + 2, \ldots, T$, there exists a Markov strategy that is optimal in period $t$.

It will follow by backward induction that there exists an optimal strategy $h^t \mapsto \psi_t(h^t)$ for every period $t = s, s + 1, \ldots, T$ that takes the Markov form $h^t \mapsto w_t \mapsto c_t^*(w_t)$.

This treats history as irrelevant, except insofar as it determines current wealth $w_t$ at the time when $c_t$ has to be chosen.
A Stochastic Difference Equation

Accordingly, suppose that the consumer pursues a Markov strategy taking the form \( w_t \mapsto c^*_t(w_t) \).

Then the Markov state variable \( w_t \) will evolve over time according to the stochastic difference equation

\[
  w_{t+1} = \phi_t(w_t, \tilde{r}_t) := \tilde{r}_t(w_t - c^*_t(w_t)).
\]

Starting at any time \( t \), conditional on initial wealth \( w_t \), this equation will have a random solution \( \tilde{w}_{t+1}^T = (\tilde{w}_{t})_{t=\tau}^{T} \) described by a unique joint conditional cdf \( F_{t+1}^T(w_{t+1}^T | w_t) \) on \( \mathbb{R}^{T-s} \).

Combined with the Markov strategy \( w_t \mapsto c^*_t(w_t) \), this generates a random consumption stream \( \tilde{c}_{t+1}^T = (\tilde{c}_{t})_{t=\tau}^{T} \) described by a unique joint conditional cdf \( G_{t+1}^T(c_{t+1}^T | w_t) \) on \( \mathbb{R}^{T-s} \).
General Finite Horizon Problem

Consider the objective of choosing $y_s$ in order to maximize

$$\mathbb{E}_s \left[ \sum_{t=s}^{T-1} u_s(x_s, y_s) + \phi_T(x_T) \right]$$

subject to the law of motion $x_{t+1} = \xi_t(x_t, y_t, \epsilon_t)$,
where the random shocks $\epsilon_t$
at different times $t = s, s + 1, s + 2, \ldots, T - 1$
are conditionally independent given $x_t, y_t$.
Here $x_T \mapsto \phi_T(x_T)$ is the terminal state valuation function.

The stochastic law of motion can also be expressed
through successive conditional probabilities $P_{t+1}(x_{t+1}|x_t, y_t)$.

The choices of $y_t$ at successive times determine
a controlled Markov process governing the stochastic transition
from each state $x_t$ to its immediate successor $x_{t+1}$. 
Backward Recurrence Relation

The optimal solution can be derived by solving the backward recurrence relation

\[
V_s(x_s) = \max_{y_s \in F_s(x_s)} \left\{ u_s(x_s, y_s) + \mathbb{E}_s \left[ V_{s+1}(x_{s+1}) \mid x_s, y_s \right] \right\}
\]

where

1. \(x_s\) denotes the “inherited state” at time \(s\);
2. \(V_s(x_s)\) is the current value in state \(x_s\) of the state value function \(X \ni x \mapsto V_s(x) \in \mathbb{R}\);
3. \(X \ni x \mapsto F_s(x) \subset Y\) is the feasible set correspondence;
4. \((x, y) \mapsto u_s(x, y)\) denotes the immediate return function in period \(s\);
5. \(X \ni x \mapsto y_s^*(x) \in F_s(x_s)\) is the optimal “strategy” or policy function;
6. The relevant terminal condition is that \(V_T(x_T)\) is given by the exogenously specified function \(\phi_T(x_T)\).
Lecture Outline

Linear Equations in One Variable

Optimal Saving

The Two Period Problem

The $T$ Period Problem

A General Problem

Infinite Time Horizon

Main Theorem

Policy Improvement

Unboundedness
An Infinite Horizon Savings Problem

Game theorists speak of the "one-shot" deviation principle. This states that if any deviation from a particular policy or strategy improves a player’s payoff, then there exists a one-shot deviation that improves the payoff.

We consider the infinite horizon extension of the consumption/investment problem already considered. Given the initial time \( s \) and initial wealth \( w_s \), this takes the form of choosing a consumption policy \( c_t(w_t) \) at times \( t = s, s + 1, s + 2, \ldots \) in order to maximize the discounted sum of total utility, given by

\[
\sum_{t=s}^{\infty} \beta^{t-s} u(c_t)
\]

subject to the accumulation equation \( w_{t+1} = \tilde{r}_t(w_t - c_t) \) as well as the inequality constraint \( w_t \geq 0 \) for \( t = s + 1, s + 2, \ldots \).
Some Assumptions

The parameter $\beta \in (0, 1)$ is the constant discount factor. Note that utility function $\mathbb{R} \ni c \mapsto u(c)$ is independent of $t$; its first two derivatives are assumed to satisfy the inequalities $u'(c) > 0$ and $u''(c) < 0$ for all $c \in \mathbb{R}_+$. The investment returns $\tilde{r}_t$ in successive periods are assumed to be i.i.d. random variables. It is assumed that $w_t$ in each period $t$ is known at time $t$, but not before.
Terminal Constraint

There has to be an additional constraint that imposes a lower bound on wealth at some time \( t \).

Otherwise there would be no optimal policy — the consumer can always gain by increasing debt (negative wealth), no matter how large existing debt may be.

In the finite horizon, there was a constraint \( w_T \geq 0 \) on terminal wealth.
But here \( T \) is effectively infinite.

One might try an alternative like

\[
\lim_{t \to \infty} \inf \beta^t w_t \geq 0
\]

But this places no limit on wealth at any finite time.

We use the alternative constraint requiring that \( w_t \geq 0 \) for all time.
The Stationary Problem

Our modified problem can be written in the following form that is independent of $s$:

$$\max_{c_0, c_1, \ldots, c_t, \ldots} \sum_{t=0}^{\infty} \beta^t u(c_t)$$

subject to the constraints $0 \leq c_t \leq w_t$ and $w_{t+1} = \tilde{r}_t (w_t - c_t)$ for all $t = 0, 1, 2, \ldots$, with $w_0 = w$, where $w$ is given.

Because the starting time $s$ is irrelevant, this is a stationary problem.

Define the state valuation function $w \mapsto V(w)$ as the maximum value of the objective, as a function of initial wealth $w$. 

University of Warwick, EC9A0 Maths for Economists

Peter J. Hammond
Bellman’s Equation

For the finite horizon problem, the principle of optimality was

\[ V_s(w_s) = \arg \max_{0 \leq c_s \leq w_s} \left\{ u_s(c_s) + \mathbb{E}_s[V_{s+1}(\tilde{r}_s(w_s - c_s))] \right\} \]

For the stationary infinite horizon problem, however, the time starting time \( s \) is irrelevant.

So the principle of optimality can be expressed as

\[ V(w) = \arg \max_{0 \leq c \leq w} \left\{ u(c) + \beta \mathbb{E}[V(\tilde{r}(w - c))] \right\} \]

The state valuation function \( w \mapsto V(w) \) appears on both left and right hand sides of this equation.

Solving it therefore involves finding a fixed point, or function, in an appropriate function space.
Isoelastic Case

We consider yet again the isoelastic case with a CES (or logarithmic) utility function that satisfies $u'(c; \epsilon) \equiv c^{-\epsilon}$ and, specifically

$$u(c; \epsilon) = \begin{cases} 
  \frac{c^{1-\epsilon}}{(1 - \epsilon)} & \text{if } \epsilon \neq 1; \\
  \ln c & \text{if } \epsilon = 1.
\end{cases}$$

Recall the corresponding finite horizon case, where we found that the solution to the corresponding equations

$$V_s(w_s) = c^*_s(w_s) = \arg \max_{0 \leq c_s \leq w_s} \left\{ u_s(c_s) + \beta E_s[V_{s+1}(\tilde{r}_s(w_s - c_s))] \right\}$$

takes the form $V_s(w) = \alpha_s + \nu_s u(w; \epsilon)$ for suitable real constants $\alpha_s$ and $\nu_s > 0$, where $\alpha_s = 0$ if $\epsilon \neq 1$. 
First-Order Condition

Accordingly, we look for a solution to the stationary problem

\[
\begin{align*}
V(w) &= c^*(w) = \arg \max_{0 \leq c \leq w} \{ u(c; \epsilon) + \beta \mathbb{E}[V(\tilde{r}(w - c))]) \}
\end{align*}
\]

taking the isoelastic form \( V(w) = \alpha + \nu u(w; \epsilon) \)
for suitable real constants \( \alpha \) and \( \nu > 0 \), where \( \alpha = 0 \) if \( \epsilon \neq 1 \).

The first-order condition for solving this concave maximization problem is

\[
c^{-\epsilon} = \beta \mathbb{E}[\tilde{r}(\tilde{r}(w - c))^{-\epsilon}] = \zeta^\epsilon(w - c)^{-\epsilon}
\]

where \( \zeta^\epsilon := \beta R^{1-\epsilon} \) with \( R \) as the certainty equivalent return defined by \( R^{1-\epsilon} := \mathbb{E}[\tilde{r}^{1-\epsilon}] \).

Hence \( c = \gamma w \) where \( \gamma^{-\epsilon} = \zeta^\epsilon (1 - \gamma)^{-\epsilon} \), implying that \( \gamma = 1/(1 + \zeta) \).
Solution in the Logarithmic Case

When \( \epsilon = 1 \) and so \( u(c; \epsilon) = \ln c \), one has

\[
V(w) = u(\gamma w; \epsilon) + \beta \{ \alpha + \nu \mathbb{E}[u(\tilde{r}(1 - \gamma)w; \epsilon)] \} \\
= \ln(\gamma w) + \beta \{ \alpha + \nu \mathbb{E}[\ln(\tilde{r}(1 - \gamma)w)] \} \\
= \ln \gamma + (1 + \beta \nu) \ln w + \beta \{ \alpha + \nu \ln(1 - \gamma) + \mathbb{E}[\ln \tilde{r}] \}
\]

This is consistent with \( V(w) = \alpha + \nu \ln w \) in case:

1. \( \nu = 1 + \beta \nu \), implying that \( \nu = (1 - \beta)^{-1} \);
2. and also \( \alpha = \ln \gamma + \beta \{ \alpha + \nu \ln(1 - \gamma) + \mathbb{E}[\ln \tilde{r}] \} \),

which implies that

\[
\alpha = (1 - \beta)^{-1} \left[ \ln \gamma + \beta \left\{ (1 - \beta)^{-1} \ln(1 - \gamma) + \mathbb{E}[\ln \tilde{r}] \right\} \right]
\]

This confirms the solution for the logarithmic case.
Solution in the CES Case

When $\epsilon \neq 1$ and so $u(c; \epsilon) = c^{1-\epsilon}/(1 - \epsilon)$, the equation

$$V(w) = u(\gamma w; \epsilon) + \beta v \mathbb{E}[u(\tilde{r}(1 - \gamma)w; \epsilon)]$$

implies that

$$(1 - \epsilon)V(w) = (\gamma w)^{1-\epsilon} + \beta v \mathbb{E}[(\tilde{r}(1 - \gamma)w)^{1-\epsilon}] = vw^{1-\epsilon}$$

where $v = \gamma^{1-\epsilon} + \beta v (1 - \gamma)^{1-\epsilon} R^{1-\epsilon}$ and so

$$v = \frac{\gamma^{1-\epsilon}}{1 - \beta (1 - \gamma)^{1-\epsilon} R^{1-\epsilon}} = \frac{\gamma^{1-\epsilon}}{1 - (1 - \gamma)^{1-\epsilon} \zeta}$$

But optimality requires $\gamma = 1/(1 + \zeta)$, implying finally that

$$v = \frac{(1 + \zeta)^{\epsilon-1}}{1 - \zeta (1 + \zeta)^{\epsilon-1}} = \frac{1}{(1 + \zeta)^{1-\epsilon} - \zeta}$$

This confirms the solution for the CES case.
Lecture Outline

Linear Equations in One Variable
Optimal Saving
The Two Period Problem
The $T$ Period Problem
A General Problem
Infinite Time Horizon

Main Theorem
Policy Improvement
Unboundedness
Uniformly Bounded Returns

Suppose that the stochastic transition from each state $x$ to the immediately succeeding state $\tilde{x}$ is specified by a conditional probability measure $B \mapsto \mathbb{P}(\tilde{x} \in B|x, u)$ on a $\sigma$-algebra of the state space.

Consider the stationary problem of choosing a policy $x \mapsto u^*(x)$ in order to maximize the infinite discounted sum of utility

$$\mathbb{E} \sum_{t=1}^{\infty} \beta^{t-1} f(x_t, u_t)$$

where $0 < \beta < 1$, with $x_1$ given and subject to $u_t \in U(x_t)$ for $t = 1, 2, \ldots$.

The return function $(x, u) \mapsto f(x, u) \in \mathbb{R}$ is uniformly bounded provided there exist a uniform lower bound $M_*$ and a uniform upper bound $M^*$ such that

$$M_* \leq f(x, u) \leq M^* \quad \text{for all} \ (x, u)$$
The Function Space

The boundedness assumption $M_* \leq f(x, u) \leq M^*$ for all $(x, u)$ ensures that, because $0 < \beta < 1$ and so $\sum_{t=1}^{\infty} \beta^{t-1} = \frac{1}{1-\beta}$, the infinite discounted sum of utility

$$W := \mathbb{E} \sum_{t=1}^{\infty} \beta^{t-1} f(x_t, u_t)$$

satisfies $(1 - \beta) W \in [M_*, M^*]$.

This makes it natural to consider the linear space $\mathcal{V}$ of all bounded functions $X \ni x \mapsto V(x) \in \mathbb{R}$ equipped with its sup norm defined by $\|V\| := \sup_{x \in X} |V(x)|$.

We will pay special attention to the subset

$$\mathcal{V}_M := \{ V \in \mathcal{V} \mid x \in X \mapsto (1 - \beta) V(x) \in [M_*, M^*] \}$$

of state valuation functions whose values $V(x)$ lie within the range of the possible values of $W$. 
Existence and Uniqueness

Theorem

Consider the Bellman equation system

\[
V(x) = \max_{u \in F(x)} \left\{ f(x, u) + \beta \mathbb{E} [V(\tilde{x}) | x, u] \right\}
\]

Under the assumption of uniformly bounded returns satisfying \( M_* \leq f(x, u) \leq M^* \) for all \((x, u)\):

1. among the set \( V_M \) of state valuation functions that satisfy the inequalities \( M_* \leq (1 - \beta) V(x) \leq M^* \) for all \( x \), there is a unique state valuation function \( x \mapsto V(x) \) that satisfies the Bellman equation system.

2. any associated policy solution \( x \mapsto u^*(x) \) determines an optimal policy that is stationary — i.e., independent of time.
Two Mappings

Given any measurable policy function $X \ni x \mapsto u(x)$ denoted by $u$, define the mapping $T^u : \mathcal{V}_M \rightarrow \mathcal{V}$ by

$$[T^u \mathcal{V}](x) := f(x, u(x)) + \beta \mathbb{E}[V(\tilde{x})|x, u(x)]$$

When the state is $x$, this gives the value $[T^u \mathcal{V}](x)$ of choosing the policy $u(x)$ for one period, and then experiencing a future discounted return $V(\tilde{x})$ after reaching each possible subsequent state $\tilde{x} \in X$.

Define also the mapping $T^* : \mathcal{V}_M \rightarrow \mathcal{V}$ by

$$[T^* \mathcal{V}](x) := \max_{u \in F(x)} \{f(x, u) + \beta \mathbb{E}[V(\tilde{x})|x, u]\}$$

These definitions allow the Bellman equation system to be rewritten as

$$V(x) = [T^* \mathcal{V}](x)$$
$$u^*(x) \in \arg \max_{u \in F(x)} [T^u \mathcal{V}](x)$$
Two Mappings of $\mathcal{V}_M$ into Itself

For all $V \in \mathcal{V}_M$, policies $u$, and $x \in X$, we have defined

$$[T^u V](x) := f(x, u(x)) + \beta \mathbb{E} [V(\tilde{x})|x, u(x)]$$

and

$$[T^* V](x) := \max_{u \in F(x)} \{ f(x, u) + \beta \mathbb{E} [V(\tilde{x})|x, u] \}$$

Recall the uniform boundedness condition $M_* \leq f(x, u) \leq M^*$, together with the assumption that $V$ belongs to the domain $\mathcal{V}_M$ of functions satisfying $M_* \leq (1 - \beta) V(\tilde{x}) \leq M^*$ for all $\tilde{x}$.

So these two definitions jointly imply that

$$[T^u V](x) \geq M_* + \beta (1 - \beta)^{-1} M_* = (1 - \beta)^{-1} M_*$$

and

$$[T^u V](x) \leq M^* + \beta (1 - \beta)^{-1} M^* = (1 - \beta)^{-1} M^*$$

Similarly, given any $V \in \mathcal{V}_M$, one has $M_* \leq (1 - \beta) [T^* V](x) \leq M^*$ for all $x \in X$.

Therefore both $V \mapsto T^u V$ and $V \mapsto T^* V$ map $\mathcal{V}_M$ into itself.
A First Contraction Mapping

The definition \( [T^u V](x) := f(x, u(x)) + \beta \mathbb{E} [V(\tilde{x})|x, u(x)] \) implies that for any two functions \( V_1, V_2 \in \mathcal{V}_M \), one has

\[
[T^u V_1](x) - [T^u V_2](x) = \beta \mathbb{E} [V_1(\tilde{x}) - V_2(\tilde{x})|x, u(x)]
\]

The definition of the sup norm therefore implies that

\[
\| T^u V_1 - T^u V_2 \| = \sup_{x \in X} | [T^u V_1](x) - [T^u V_2](x) | \\
= \sup_{x \in X} \left| \beta \mathbb{E} [V_1(\tilde{x}) - V_2(\tilde{x})|x, u(x)] \right| \\
= \beta \sup_{x \in X} \left| \mathbb{E} [V_1(\tilde{x}) - V_2(\tilde{x})|x, u(x)] \right| \\
\leq \beta \sup_{\tilde{x} \in X} | V_1(\tilde{x}) - V_2(\tilde{x}) | \\
= \beta \| V_1 - V_2 \|
\]

Hence \( V \mapsto T^u V \) is a contraction mapping with factor \( \beta < 1 \) that maps the bounded subset \( \mathcal{V}_M \) of the normed linear space \( \mathcal{V} \) into itself.
For each fixed policy $u$, the contraction mapping $V \mapsto T^u V$ mapping the space $\mathcal{V}_M$ into itself has a unique fixed point in the form of a function $V^u \in \mathcal{V}_M$.

Furthermore, given any initial function $V \in \mathcal{V}_M$, consider the infinite sequence of mappings $[T^u]^k V$ $(k \in \mathbb{N})$ that result from applying the operator $T^u$ iteratively $k$ times.

The contraction mapping property of $T^u$ implies that $\| [T^u]^k V - V^u \| \to 0$ as $k \to \infty$. 
Characterizing the Fixed Point, I

Starting from \( V_0 = 0 \) and given any initial state \( x \in X \), note that

\[
[T^u]^k V_0(x) = \left[ T^u \right] \left( [T^u]^{k-1} V_0 \right)(x) \\
= f(x, u(x)) + \beta E \left[ ([T^u]^{k-1} V_0)(\tilde{x})|x, u(x) \right]
\]

It follows by induction on \( k \) that \( [T^u]^k V_0(\bar{x}) \) equals the expected discounted total payoff \( \mathbb{E} \sum_{t=1}^{k} \beta^{t-1} f(x_t, u_t) \)
of starting from \( x_0 = \bar{x} \)
and then following the policy \( x \mapsto u(x) \) for \( k \) subsequent periods.

Taking the limit as \( k \to \infty \), it follows that for any state \( \bar{x} \in X \),
the value \( V^u(\bar{x}) \) of the fixed point in \( V_M \)
is the expected discounted total payoff

\[
\mathbb{E} \sum_{t=1}^{\infty} \beta^{t-1} f(x_t, u_t)
\]
of starting from \( x_0 = \bar{x} \)
and then following the policy \( x \mapsto u(x) \) for ever thereafter.
A Second Contraction Mapping

Recall the definition

\[ [T^* V](x) := \max_{u \in F(x)} \{ f(x, u) + \beta \mathbb{E} [V(\tilde{x})|x, u] \} \]

Given any state \( x \in X \) and any two functions \( V_1, V_2 \in \mathcal{V}_M \), define \( u_1, u_2 \in F(x) \) so that for \( k = 1, 2 \) one has

\[ [T^* V_k](x) = f(x, u_k) + \beta \mathbb{E} [V_k(\tilde{x})|x, u_k] \]

Note that \( [T^* V_2](x) \geq f(x, u_1) + \beta \mathbb{E} [V_2(\tilde{x})|x, u_1] \) implying that

\[ [T^* V_1](x) - [T^* V_2](x) \leq \beta \mathbb{E} [V_1(\tilde{x}) - V_2(\tilde{x})|x, u_1] \]
\[ \leq \beta \| V_1 - V_2 \| \]

Similarly, interchanging 1 and 2 in the above argument gives \( [T^* V_2](x) - [T^* V_1](x) \leq \beta \| V_1 - V_2 \| \).

Hence \( \| T^* V_1 - T^* V_2 \| \leq \beta \| V_1 - V_2 \| \), so \( T^* \) is also a contraction.
Applying the Contraction Mapping Theorem, II

Similarly the contraction mapping $V \mapsto T^* V$
has a unique fixed point in the form of a function $V^* \in \mathcal{V}_M$
such that $V^*(\bar{x})$ is the maximized expected discounted total payoff
of starting in state $x_0 = \bar{x}$
and following an optimal policy for ever thereafter.

Moreover, $V^* = T^* V^* = T^{u^*} V$.

This implies that $V^*$ is also the value
of following the policy $x \mapsto u^*(x)$ throughout,
which must therefore be an optimal policy.
Characterizing the Fixed Point, II

Starting from $V_0 = 0$ and given any initial state $x \in X$, note that

$$\left[T^*_k \right] V_0(x) = \left[T^*_k \right] \left( \left[T^*_k \right]^{k-1} V_0 \right) (x)$$

$$= \max_{u \in F(x)} \{ f(x, u) + \beta \mathbb{E} \left[ \left( \left[T^*_k \right]^{k-1} V_0 \right) (\tilde{x}) | x, u \right] \}$$

It follows by induction on $k$ that $\left[T^*_k \right] V_0(\tilde{x})$ equals the maximum possible expected discounted total payoff $\mathbb{E} \sum_{t=1}^{k} \beta^{t-1} f(x_t, u_t)$ of starting from $x_1 = \tilde{x}$ and then following the “backward” sequence of optimal policies $(u_1^*, u_2^*, \ldots, u_{k-2}^*, u_{k-1}^*, u_k^*)$, where for each $k$ the policy $x \mapsto u_k^*(x)$ is optimal when $k$ periods remain.
Method of Successive Approximation

The method of successive approximation starts with an arbitrary function \( V_0 \in \mathcal{V}_M \).

For \( k = 1, 2, \ldots, \), it then repeatedly solves the pair of equations

\[
V_k = T^* V_{k-1} = Tu_k^* V_{k-1}
\]

to construct sequences of:

1. state valuation functions \( X \ni x \mapsto V_k(x) \in \mathbb{R} \);
2. policies \( X \ni x \mapsto u_k^*(x) \in F(x) \) that are optimal given that one applies the preceding state valuation function \( X \ni \tilde{x} \mapsto V_{k-1}(\tilde{x}) \in \mathbb{R} \) to each immediately succeeding state \( \tilde{x} \).

Because the operator \( V \mapsto T^* V \) on \( \mathcal{V}_M \) is a contraction mapping, the method produces a convergent sequence \( (V_k)_{k=1}^\infty \) of state valuation functions whose limit satisfies

\[
V^* = T^* V^* = Tu^* V^*
\]

for a suitable policy \( X \ni x \mapsto u^*(x) \in F(x) \).
Lecture Outline

Linear Equations in One Variable

Optimal Saving

The Two Period Problem

The $T$ Period Problem

A General Problem

Infinite Time Horizon

Main Theorem

Policy Improvement

Unboundedness
Monotonicity

For all functions $V \in \mathcal{V}_M$, policies $u$, and states $x \in X$, we have defined

$$[T^u V](x) := f(x, u(x)) + \beta \mathbb{E} [V(\tilde{x})|x, u(x)]$$

and

$$[T^* V](x) := \max_{u \in F(x)} \{f(x, u) + \beta \mathbb{E} [V(\tilde{x})|x, u]\}$$

Notation

Given any pair $V_1, V_2 \in \mathcal{V}_M$, we write $V_1 \geq V_2$ to indicate that the inequality $V_1(x) \geq V_2(x)$ holds for all $x \in X$.

Definition

An operator $\mathcal{V}_M \ni V \mapsto TV \in \mathcal{V}_M$ is monotone just in case whenever $V_1, V_2 \in \mathcal{V}_M$ satisfy $V_1 \geq V_2$, one has $TV_1 \geq TV_2$.

Theorem

The following operators on $\mathcal{V}_M$ are monotone:

1. $V \mapsto T^u V$ for all policies $u$;
2. $V \mapsto T^* V$ for the optimal policy.
Proof that $T^u$ is Monotone

Given any state $x \in X$ and any two functions $V_1, V_2 \in \mathcal{V}_M$, the definition of $T^u$ implies that

$$[T^u V_1](x) := f(x, u(x)) + \beta \mathbb{E} [V_1(\tilde{x})|x, u(x)]$$

and

$$[T^u V_2](x) := f(x, u(x)) + \beta \mathbb{E} [V_2(\tilde{x})|x, u(x)]$$

Subtracting the second equation from the first implies that

$$[T^u V_1](x) - [T^u V_2](x) = \beta \mathbb{E} [V_1(\tilde{x}) - V_2(\tilde{x})|x, u(x)]$$

If $V_1 \geq V_2$ and so the inequality $V_1(\tilde{x}) \geq V_2(\tilde{x})$ holds for all $\tilde{x} \in X$, it follows that $[T^u V_1](x) \geq [T^u V_2](x)$.

Since this holds for all $x \in X$, we have proved that $T^u V_1 \geq T^u V_2$. \qed
Proof that $T^*$ is Monotone

Given any state $x \in X$ and any two functions $V_1, V_2 \in \mathcal{V}_M$, define $u_1, u_2 \in F(x)$ so that for $k = 1, 2$ one has

$$[T^* V_k](x) = \max_{u \in F(x)} \{f(x, u) + \beta \mathbb{E} [V_k(\tilde{x})|x, u]\}$$

$$= [T^{u_k} V_k](x) = f(x, u_k) + \beta \mathbb{E} [V_k(\tilde{x})|x, u_k]$$

It follows that

$$[T^* V_1](x) \geq f(x, u_2) + \beta \mathbb{E} [V_1(\tilde{x})|x, u_2]$$

and

$$[T^* V_2](x) = f(x, u_2) + \beta \mathbb{E} [V_2(\tilde{x})|x, u_2]$$

Subtracting the second equation from the first inequality gives

$$[T^* V_1](x) - [T^* V_2](x) \geq \beta \mathbb{E} [V_1(\tilde{x}) - V_2(\tilde{x})|x, u_2]$$

If $V_1 \geq V_2$ and so the inequality $V_1(\tilde{x}) \geq V_2(\tilde{x})$ holds for all $\tilde{x} \in X$, it follows that $[T^* V_1](x) \geq [T^* V_2](x)$.

Since this holds for all $x \in X$, we have proved that $T^* V_1 \geq T^* V_2$. 

\(\square\)
The method of policy improvement starts with any fixed policy $u_0$ or $X \ni x \mapsto u_0(x) \in F(x)$, along with the value $V^{u_0} \in \mathcal{V}_M$ of following that policy for ever.

The value $V^{u_0}$ is the unique fixed point satisfying $V^{u_0} = T^{u_0} V^{u_0}$ which belongs to the domain $\mathcal{V}_M$ of suitably bounded functions.

At each step $k = 1, 2, \ldots$, given the previous policy $u_{k-1}$ and associated value $V^{u_{k-1}}$ satisfying $V^{u_{k-1}} = T^{u_{k-1}} V^{u_{k-1}}$:

1. the policy $u_k$ is chosen so that $T^* V^{u_{k-1}} = T^{u_k} V^{u_{k-1}}$;
2. the state valuation function $x \mapsto V_k(x)$ is chosen as the unique fixed point in $\mathcal{V}_M$ of the operator $T^{u_k}$. 
Theorem

The infinite sequence \((u_k, V^{u_k})_{k \in \mathbb{N}}\)
consisting of pairs of policies \(u_k\)
with their associated valuation functions \(V^{u_k} \in \mathcal{V}_M\) satisfies

1. \(V^{u_k} \geq V^{u_{k-1}}\) for all \(k \in \mathbb{N}\) (policy improvement);
2. \(\|V^{u_k} - V^*\| \to 0\) as \(k \to \infty\),
   where \(V^*\) is the infinite-horizon optimal
   state valuation function in \(\mathcal{V}_M\) that satisfies \(T^*V^* = V^*\).
Proof of Policy Improvement

By definition of the optimality operator $T^*$, one has $T^* V \geq T^u V$ for all functions $V \in \mathcal{V}_M$ and all policies $u$.

So at each step $k$ of the policy improvement routine, one has

$$T^{u_k} V^{u_{k-1}} = T^* V^{u_{k-1}} \geq T^{u_{k-1}} V^{u_{k-1}} = V^{u_{k-1}}$$

In particular, $T^{u_k} V^{u_{k-1}} \geq V^{u_{k-1}}$.

Now, applying successive iterations of the monotonic operator $T^{u_k}$ implies that

$$V^{u_{k-1}} \leq T^{u_k} V^{u_{k-1}} \leq [T^{u_k}]^2 V^{u_{k-1}} \leq \ldots$$

$$\ldots \leq [T^{u_k}]^r V^{u_{k-1}} \leq [T^{u_k}]^{r+1} V^{u_{k-1}} \leq \ldots$$

But the definition of $V^{u_k}$ implies that for all $V \in \mathcal{V}_M$, including $V = V^{u_{k-1}}$, one has $\| [T^{u_k}]^r V - V^{u_k} \| \to 0$ as $r \to \infty$.

Hence $V^{u_k} = \sup_r [T^{u_k}]^r V^{u_{k-1}} \geq V^{u_{k-1}}$, thus confirming that the policy $u_k$ does improve $u_{k-1}$. 

\[\square\]
Proof of Convergence

Recall that at each step $k$ of the policy improvement routine, one has $T^u_k V^{u_k-1} = T^* V^{u_k-1}$ and also $T^u_k V^{u_k} = V^{u_k}$.

Now, for each state $x \in X$, define $\hat{V}(x) := \sup_{k \in \mathbb{N}} V^{u_k}(x)$.

Because $V^{u_k} \geq V^{u_k-1}$ and $T^u_k$ is monotonic, one has $V^{u_k} = T^u_k V^{u_k} \geq T^u_k V^{u_k-1} = T^* V^{u_k-1}$.

Next, because $T^*$ is monotonic, it follows that

$$\hat{V} = \sup_k V^{u_k} \geq \sup_k T^* V^{u_k-1} = T^*(\sup_k V^{u_k-1}) = T^* \hat{V}$$

Similarly, monotonicity of $T^*$ and the definition of $T^*$ imply that

$$\hat{V} = \sup_k V^{u_k} = \sup_k T^u_k V^{u_k} \leq \sup_k T^* V^{u_k} = T^*(\sup_k V^{u_k}) = T^* \hat{V}$$

Hence $\hat{V} = T^* \hat{V} = V^*$, because $T^*$ has a unique fixed point.

Therefore $V^* = \sup_k V^{u_k}$ and so, because the sequence $V^{u_k}(x)$ is non-decreasing, one has $V^{u_k}(x) \to V^*(x)$ for each $x \in X$. □
Lecture Outline

Linear Equations in One Variable
Optimal Saving
The Two Period Problem
The $T$ Period Problem
A General Problem
Infinite Time Horizon
Main Theorem
Policy Improvement
Unboundedness
Unbounded Utility

In economics the boundedness condition $M_* \leq f(x, u) \leq M^*$ is rarely satisfied!

Consider for example the isoelastic utility function

$$u(c; \epsilon) = \begin{cases} 
\frac{c^{1-\epsilon}}{1 - \epsilon} & \text{if } \epsilon > 0 \text{ and } \epsilon \neq 1 \\
\ln c & \text{if } \epsilon = 1 
\end{cases}$$

This function is obviously:

1. bounded below but unbounded above in case $0 < \epsilon < 1$;
2. unbounded both above and below in case $\epsilon = 1$;
3. bounded above but unbounded below in case $\epsilon > 1$.

Also commonly used is the negative exponential utility function defined by $u(c) = -e^{-\alpha c}$ where $\alpha$ is the constant absolute rate of risk aversion (CARA).

This function is bounded above and, provided that $c \geq 0$, also below.
Warning Example: Statement of Problem

The following example shows that there can be irrelevant **unbounded** solutions to the Bellman equation.

**Example**
Consider the problem of maximizing $\sum_{t=0}^{\infty} \beta^t (1 - u_t)$

where $u_t \in [0, 1]$, $0 < \beta < 1$, and $x_{t+1} = \frac{1}{\beta} (x_t + u_t)$, with $x_0 > 0$.

Notice that $x_{t+1} \geq \frac{1}{\beta} x_t$ implying that $x_t \geq \beta^{-t} x_0 \to \infty$ as $t \to \infty$.

Of course the return function $[0, 1] \ni u \mapsto f(x, u) = 1 - u \in [0, 1]$ is uniformly bounded.
Warning Example: Unbounded Spurious Solution

The Bellman equation is

\[
\begin{align*}
J(x) &= u^*(x) = \arg \max_{u \in [0,1]} \left\{ 1 - u + \beta J \left( \frac{1}{\beta} (x + u) \right) \right\} \\
\end{align*}
\]

Even though the return function is uniformly bounded, this Bellman equation has an unbounded spurious solution.

Indeed, we find a spurious solution with \( J(x) \equiv \gamma + x \) for a suitable constant \( \gamma \).

The condition for this to solve the Bellman equation is that

\[
\begin{align*}
\gamma + x &= \max_{u \in [0,1]} \left\{ 1 - u + \beta \left[ \gamma + \frac{1}{\beta} (x + u) \right] \right\} \\
&= \max_{u \in [0,1]} \{ 1 + \beta \gamma + x \} = 1 + \beta \gamma + x
\end{align*}
\]

which is true iff \( \gamma = 1 + \beta \gamma \) and so \( \gamma = (1 - \beta)^{-1} \).
Warning Example: True Solution

The problem is to maximize \( \sum_{t=0}^{\infty} \beta^t (1 - u_t) \)

where \( u_t \in [0, 1] \), \( 0 < \beta < 1 \), and \( x_{t+1} = \frac{1}{\beta} (x_t + u_t) \), with \( x_0 > 0 \).

The obvious optimal policy is to choose \( u_t = 0 \) for all \( t \),
giving the maximized value \( J(x) = \sum_{t=0}^{\infty} \beta^t = (1 - \beta)^{-1} \).

Indeed the bounded function \( J(x) = (1 - \beta)^{-1} \),
together with \( u^* = 0 \), both independent of \( x \),
do indeed solve the Bellman equation

\[
J(x) = \max_{u \in [0,1]} \left\{ 1 - u + \beta J \left( \frac{1}{\beta} (x + u) \right) \right\}
\]

\[
= \max_{u \in [0,1]} \left\{ 1 - u + \beta (1 - \beta)^{-1} \right\}
\]

\[
= 1 + \frac{\beta}{1 - \beta} = \frac{1}{1 - \beta}
\]