Economics 600: Mathematical Economics
August 2004

Matthew Chesnes

Updated: August 20, 2004

1 Preliminary Concepts

1.1 Some Examples

1.2 Continuity and Linearity

- For a vector $x \in \mathbb{R}^n$, the length of the vector, or its NORM, is given by:

  $$||x|| = \sqrt{x_1^2 + x_2^2 + \cdots + x_n^2}.$$

  This is also known as the Euclidean norm.

- **Definition:** A NORM is any real valued function, $|| \cdot || : A \mapsto \mathbb{R}$ that satisfies:
  - $||x|| \geq 0 \; \forall \; x$ and $||x|| = 0$ iff $x = 0$.
  - $||x + y|| \leq ||x|| + ||y|| \; \forall \; x, y \in A$. (Triangle Inequality).
  - $||ax|| = |a| * ||x|| \; \forall \; a \in \mathbb{R}$ and $x \in A$.

- **Definition:** A sequence of elements CONVERGES to a point x in $\mathbb{R}^n, \{x_n\}_{n=1}^{\infty}$, if for every $\delta > 0$, there is a number, N, s.t. $\forall \; n \geq N, ||x_n - x|| < \delta$.

- **Definition:** A function, $f : \mathbb{R}^n \mapsto \mathbb{R}^m$ is CONTINUOUS at a point $x \in \mathbb{R}^n$ if for all sequences $\{x_n\}$ converging to x, the sequence $\{f(x_n)\}$ converges to $f(x)$. A function is continuous if it is continuous at all points in its domain.

- **Definition:** A function $f : V \mapsto W$ is LINEAR if for any two real numbers, a and b and two elements x and y in V, then,

  $$f(ax + by) = af(x) + bf(y).$$

- **Definition:** A function $f : \mathbb{R} \mapsto \mathbb{R}$ given by $f(x) = mx + b$ is an AFFINE function. (Not linear, does not go through the origin).

- **Definition:** A function f is UPPER SEMI-CONINUOUS if $\forall \; x_n \longrightarrow x, lim_{x \to \infty} f(x_n) \leq f(x)$.

1.3 Vector Geometry

- **Definition:** For 2 vectors, $x, y \in \mathbb{R}^n$, the INNER PRODUCT of x and y is given by:

  $$x \cdot y = x_1 y_1 + x_2 y_2 + \cdots + x_n y_n = x'y.$$

- **Theorem:** (SB 10.4) Let V and $W \in \mathbb{R}^n$. Then:
  - if $V'W > 0$, V and W form an acute angle with each other.
  - if $V'W < 0$, V and W form an obtuse angle with each other.
  - if $V'W = 0$, V and W are orthogonal (perpendicular) to each other.

1.4 Hyperplane

- **Definition:** A HYPERPLANE is the set of points $\{x : f(x) = c\}$ where f is a linear function and $c \in \mathbb{R}$. So a line in $\mathbb{R}^2$ or a plane in $\mathbb{R}^3$ are both hyperplanes.

- **Definition:** A HALF SPACE is the set of points on one side or the other of a hyperplane:
$$HS(f) = \{x : f(x) \leq c\} \quad HS(f) = \{x : f(x) \geq c\}.$$

- **Definition:** A hyperplane SEPARATES 2 sets C_1 and C_2 if $\forall\, x \in C_1, f(x) \leq c$ and $\forall\, x \in C_2, f(x) \geq c$.

- **Definition:** If C lies in a half space defined by H and H contains a point on the boundry of C, then we say the H is a SUPPORTING hyperplane of C.

- **Definition:** A set $C \in \mathbb{R}^n$ is convex if $\forall\, x$ and $y \in C$ and $\forall\, \alpha \in [0,1]$,
$$\underbrace{\alpha x + (1-\alpha)y \in C}_{ConvexCombination}.$$

- **Theorem:** (Takayama pp39-49). Suppose x and y are non-empty convex sets in $\mathbb{R}^n$ s.t. the interior of $y \cap x = \emptyset$ and the interior of y is $\neq \emptyset$, then $\exists$ a vector $a \in \mathbb{R}^n$ which is the defining vector of a separating hyperplane between x and y. That is, $\forall\, x \in X, a'x \leq c$ and $\forall\, y \in Y, a'y \geq c$. Note this is a sufficient condition. We could still have a separating hyperplane between a concave and a convex set.

- **Definition:** The graph of a function $f : \mathbb{R}^n \mapsto \mathbb{R}$ is a subset in $\mathbb{R}^{n+1}$ given by
$$\{(x,y) : x \in \mathbb{R}^n \;\; and \;\; y = f(x)\}.$$
A function may go from $\mathbb{R}$ to $\mathbb{R}$ but the graph of the function lies in $\mathbb{R}^2$.

1.5 Derivatives and Gradients

- **Definition:** Let $f : \mathbb{R} \mapsto \mathbb{R}$. The DERIVATIVE of the function f at some point x in the domain is :
$$f'(x) = lim_{h \to 0} \frac{f(x+h) - f(x)}{h}.$$
IF IT EXISTS! Note that we can approximate the value of f at a point close to x using the derivative:
$$f(x+h) \approx f(x) + f'(x)h.$$

- **Definition:** Let $f : \mathbb{R}^n \mapsto \mathbb{R}$. The PARTIAL DERIVATIVE of f with respect to its i^{th} component is:
$$\frac{\partial f(x)}{\partial x_i} = lim_{h \to 0} \frac{f(x_i + h; x_{-i}) - f(x_i; x_{-i})}{h}.$$

- **Definition:** The GRADIENT of f at the point x is :

$$\nabla f(x) = [\frac{\partial f(x)}{\partial x_1}, \frac{\partial f(x)}{\partial x_2}, \cdots, \frac{\partial f(x)}{\partial x_n}]'.$$

  This is also called the Jacobian. Notice it is an $nx1$ vector.

- **Definition:** The DERIVATIVE of a function f at x is:

$$df = [\nabla f(x)]' = [\frac{\partial f(x)}{\partial x_1}, \frac{\partial f(x)}{\partial x_2}, \cdots, \frac{\partial f(x)}{\partial x_n}].$$

  Notice it is a $1xn$ row vector and is the transpose of the jacobian or gradient.

- Thus following the logic from before, we can approximate a small move of v from the point x using the derivative (gradient) s.t.:

$$f(x+v) \approx f(x) + v \cdot \nabla f(x) = f(x) + v'\nabla f(x).$$

  Note that x and v are both $nx1$ and v is "small."

- Let $f : \mathbb{R}^n \mapsto \mathbb{R}$ and is C^2 (twice continuously differentiable). Let $f_{ij} = \frac{\partial^2 f}{\partial x_i \partial x_j}$. Define the HESSIAN of the function f as:

$$H = D^2 f(x) = \begin{bmatrix} f_{11} & f_{12} & \dots & f_{1n} \\ f_{21} & f_{22} & \dots & f_{2n} \\ \dots & \dots & \dots & \dots \\ f_{n1} & f_{n2} & \dots & f_{nn} \end{bmatrix}. \quad (1)$$

- **Theorem:** (Young's Theorem, SB 14.5) If $f : \mathbb{R}^n \mapsto \mathbb{R}$ and is C^2, then $f_{ij}(x) = f_{ji}(x)$. Thus the order that you take the partial derivatives does not matter.

1.6 Homogeneous and Homothetic Functions

- **Definition:** A function $f : \mathbb{R}^n \mapsto \mathbb{R}$ is HOMOGENEOUS of degree k if :

$$f(tx_1, tx_2, \cdots, tx_n) = t^k f(x_1, x_2, \cdots, x_n) \ \forall \ t > 0.$$

  The special case when$f$ is homogeneous of order 1, ie $f(tx) = tf(x)$, we say that f is linearly homogeneous.

- **Theorem:** (Euler's Theorem, SB 20.4). If f is C^1 and is also homogeneous of degree k, then,

$$x \cdot \nabla f(x) = kf(x).$$

  Proof of this is in the written notes.

- **Definition:** A RAY through some point x in $\mathbb{R}^n$ is the set $\{x' \in \mathbb{R}^n : x' = tx, t \in \mathbb{R}\}$. The gradient of a homogeneous function is essentially the same along any ray.

- **Definition:** A function $f : \mathbb{R}^n \mapsto \mathbb{R}$ is HOMOTHETIC if:

$$f(x) = h(v(x)),$$

Where $h : \mathbb{R} \mapsto \mathbb{R}$ is strictly increasing and $v : \mathbb{R}^n \mapsto \mathbb{R}$ is homogeneous of degree k. Thus f must be a monotonic transformation of a homogeneous function. Thus, homogeneity implies homotheticity but homotheticity does NOT imply homogeneity.

1.7 More Geometry of vectors in $\mathbb{R}^n$

- **Definition:** Let $f : \mathbb{R}^n \mapsto \mathbb{R}$. The LEVEL SET is $\{x : f(x) = c\}$, where x is in the domain of f and $c \in \mathbb{R}$. The set, $\{x : f(x) \geq c\}$ is the upper contour set.

- **Theorem:** (SB 14.2) Consider a C^1 function $f : \mathbb{R}^n \mapsto \mathbb{R}$. the gradient of $f, \nabla f(x)$, points in the DIRECTION OF GREATEST INCREASE of f moving from x. Along a level set, $df = 0$, by definition. Thus, $v'\nabla f(x) = 0$, which means that v is orthogonal to the gradient. See graph in notes. Note that $v'\nabla f(x)$ is "how much f changes when we move in the direction of v."

2 Concepts and Problems in Unconstrained Optimization

2.1 Convexity, Concavity, and Quasi-Concavity

- **Definition:** A function $f : \mathbb{R}^n \mapsto \mathbb{R}$ is CONCAVE if $\forall\ x, y$ in the domain of f and $\forall\ t \in [0, 1]$,

$$f(tx + (1-t)y) \geq tf(x) + (1-t)f(y).$$

- **Definition:** A function $f : \mathbb{R}^n \mapsto \mathbb{R}$ is STRICTLY CONCAVE if $\forall\ x, y$ in the domain of f and $\forall\ t \in [0, 1]$,

$$f(tx + (1-t)y) > tf(x) + (1-t)f(y).$$

- **Definition:** A function $f : \mathbb{R}^n \mapsto \mathbb{R}$ is CONVEX if $\forall\ x, y$ in the domain of f and $\forall\ t \in [0, 1]$,

$$f(tx + (1-t)y) \leq tf(x) + (1-t)f(y).$$

Or a function is convex is $-f$ is concave.

- **Definition:** A function $f : \mathbb{R}^n \mapsto \mathbb{R}$ is STRICTLY CONVEX if $\forall\ x, y$ in the domain of f and $\forall\ t \in [0, 1]$,

$$f(tx + (1-t)y) < tf(x) + (1-t)f(y).$$

- **Note:** A concave function outlines a convex set in $\mathbb{R}^{n+1}$. That is, f is concave iff $\{(x, y) : y \leq f(x)\}$ is convex.

- **Fact:** If $f : \mathbb{R} \mapsto \mathbb{R}$ is C^2, f is concave iff:

$$f''(x) \leq 0 \; \forall \; x.$$

So the first derivative is decreasing when we have a concave function (ie, we'll be looking for a max).

- **Definition:** Another definition of CONCAVE. Suppose $f : \mathbb{R}^n \mapsto \mathbb{R}$ is C^1. Then f is concave iff for any two points x and y,

$$f(y) \leq f(x) + \nabla f(x)'(y - x).$$

In other words, the graph of f must lie below the supporting hyperplane.

- **Definition:** Let A be an nxn symmetric matrix. Then we say that this matrix, A, is :
  - Negative Semi-Definite if $x'Ax \leq 0 \; \forall \; x$.
  - Negative Definite if $x'Ax < 0 \; \forall \; x$.
  - Positive Semi-Definite if $x'Ax \geq 0 \; \forall \; x$.
  - Positive Definite if $x'Ax > 0 \; \forall \; x$. Note that $x'Ax$ is a quadratic form.

- **Definition:** Let A be an nxn matrix. The k^{th} order LEADING PRINCIPAL MINOR (l.p.m.) of A is the determinate of the kxk matrix obtained by deleting the LAST $n-k$ rows and columns of A.

- **Definition:** The k^{th} order PRINCIPAL MINOR (p.m.) of A is the determinant of the matrix obtained by deleting ANY $n - k$ rows and the same $n - k$ columns. There are always $2^n - 1$ p.m's.

- **Fact:** An nxn symmetric matrix, A, is:
  - Positive definite iff its n l.p.m. are strictly postive.
  - Negative definite iff its n l.p.m. alternate in sign starting with the first l.p.m., a_{11} being negative.
  - Postive Semi-Definite iff all its p.m's. non-negative (≥ 0).
  - Negative Semi-Definite iff all its p.m's. of ODD order are ≤ 0 and all its p.m's of EVEN order are ≥ 0.

- **Theorem:** (SB 21.5) Suppose that $f : \mathbb{R}^n \mapsto \mathbb{R}$ is C^2. Then f is concave iff the hessian, $D^2 f(x)$, is NEGATIVE SEMI-DEFINITE $\forall \; x$. f is convex iff the hessian is POSITIVE SEMI-DEFINITE $\forall \; x$. So we have the sufficient condition that if the hessian is negative definite, the function is strictly concave. If it's positive definite, the function is strictly convex.

- **Claim:** Suppose a function $f : \mathbb{R}^n \mapsto \mathbb{R}$ is concave. Then the set of points in $\mathbb{R}^n$ given by:

$$\{x : f(x) - b \geq 0\},$$

is a convex set. Recall that a hyperplane (x such that $a'x = b$) defines two half-spaces. Half spaces are always convex and the intersection of half spaces are also convex.

- **Definition:** A function $f : \mathbb{R}^n \mapsto \mathbb{R}$ is QUASI-CONCAVE if for all $c \in \Re$, the set $\{x \in \Re^n : f(x) \geq c\}$ is convex. (The upper contour sets are convex sets). Alternatively, a function is quasi-concave if:

$$f(tx + (1-t)y) \geq min\{f(x), f(y)\} \; \forall \; x, y \in \Re^n, t \in [0,1].$$

Also, a function is quasi-convex if:

$$f(tx + (1-t)y) \leq max\{f(x), f(y)\} \; \forall \; x, y \in \Re^n, t \in [0,1].$$

And of course, a function is STRICTLY quasi-convex (concave) is the equalities are strict.

- **Claim:** If f is concave, it is quasi-concave. However, the opposite is NOT true. Also, any monotonic transformation of a concave function is quasi-concave.

- Suppose $f : \mathbb{R}^n \mapsto \mathbb{R}$ is C^2. Then the BORDERED HESSIAN is defined as:

$$BH = \begin{bmatrix} 0 & f_1 & \dots & f_n \\ f_1 & f_{11} & \dots & f_{1n} \\ f_2 & f_{21} & \dots & f_{2n} \\ \dots & \dots & \dots & \dots \\ f_n & f_{n1} & \dots & f_{nn} \end{bmatrix}. \tag{2}$$

Note the gradient and its transpose are positioned along the left and top of this matrix. The $(1,1)$ element is always 0.

- **Theorem:** (SB 21.19) This is a sufficient condition for quasi-concavity. Calculate the l.p.m. of the BH of f starting from $k = 3$ onward. If they alternate in sign, with the first being positive $\forall \; x$, then f is quasi-concave. If they are ALL negative, f is quasi-convex. Note the first l.p.m. is 0 and the second is always positive.

2.2 Necessary Conditions for an Optimum

- **Definition:** A point $x^* \in \Re^n$ is a global maximum of a function f if $\forall \; x \in \Re^n, f(x^*) \geq f(x)$. A point $x^* \in \Re^n$ is a strict global max of f if $\forall \; x \in \Re^n \neq x^*, f(x^*) > f(x)$.

- **Definition:** Let $\epsilon > 0$. A set $B_\epsilon(x) = \{y : ||y - x|| < \epsilon\}$ is an open ball around x.

- **Definition:** A point $x^* \in \Re^n$ is a local maximum of a function f, if $\exists$ an open ball around $x^*, B_\epsilon(x^*)$ s.t. $\forall\, x \in B_\epsilon(x^*)$,

$$f(x^*) \geq f(x).$$

Thus a global max is clearly also a local max but not necessarily vice versa.

- NOTE 1. A is necessary for B means that: $B \longrightarrow A$ but A does not directly imply B. A just has to happen if B is happening.

- NOTE 2. A is sufficient for B means that $A \longrightarrow B$ but B does not directly imply A. A is enough to be sure B is happening but A could also not happen, and still have B happening.

- Recall, $f(x+v) \approx f(x) + v'\nabla f(x)$. Thus **Theorem:** Suppose f is C^1. If x^* is a local maximum of f, then $\nabla f(x^*) = 0$. This is a NECESSARY condition. See notes for a non-rigerous proof.

- **Theorem:** If x^* is a local max of f and f is continuous, then $\exists$ an open ball around x^*, $B_\epsilon(x^*)$ s.t. f is concave on $B_\epsilon(x^*)$.

- **Corollary:** Suppose f is C^2. If x^* is a local max of f, then $D^2 f(x^*)$ is negative semi-definite. This is a NECESSARY condition! Local max $\rightarrow$ $D^2 f(x)$ is locally concave.

Non-Differential functions

- **Definition:** The SUPER-DIFFERENTIAL of a concave function, f, at a point, x^*, is the set of all supporting hyperplanes of the graph of f at the point $x^*, f(x^*)$. If f is convex, we call this the SUB-DIFFERENTIAL.

- **Definition:** The SUPER-GRADIENT of a function, f, at a point, x^*, is any element (a hyperplane) of the Super Differential of f at x^*.

- **Theorem:** If x^* is an unconstrained local maximum of a function $f : \mathbb{R}^n \mapsto \mathbb{R}$, then one Super-Gradient at x^* has slope ZERO. Note that if the function is differential, then the Super Differential and the Gradient are the same thing. There is only ONE supporting hyperplane for a differential function, while there may be many for a non-differential function. See graph in notes.

2.3 Sufficient Conditions for an Optimum

- **Theorem:** If a function, f, is concave ($f : \mathbb{R}^n \mapsto \mathbb{R}$) and C^1, then:

$$\nabla f(x^*) = 0 \Longrightarrow x^* \text{is a global max of } f.$$

$\nabla f(x^*) = 0$ is both necessary and sufficient for a max of f at x^*. Note that if f is strictly concave, $\nabla f(x^*) = 0$ implies that x^* is a strict global maximum.

2.4 Minimizing Versus Maximizing

- **Theorem:** Suppose $f : \mathbb{R}^n \mapsto \mathbb{R}$. Then x^* solves $max\ \{f(x)\}$ iff x^* also solves $min\ \{-f(x)\}$.
- Properties of Unconstrained Maximization. Suppose $\nabla f(x^*) = 0$ (Necessary for an optimum).
  - If $D^2 f(x^*) = H(x^*)$ is Negative Definite, then x^* is a strict local max. (Sufficient Condition)
  - If $D^2 f(x^*) = H(x^*)$ is Positive Definite, then x^* is a strict local min. (Sufficient Condition)
  - If $D^2 f(x^*) = H(x^*)$ is Indefinite, then x^* is not a local max or min. (Sufficient Condition)
  - If x^* is a local max, $H(x^*)$ is Negative Semi-Definite. (Necessary Condition)
  - If x^* is a local min, $H(x^*)$ is Positive Semi-Definite. (Necessary Condition)
- Generalizing to $\forall\ x$.
  - If $\forall\ x, H(x)$ is Negative Definite, then x^* is a strict global maximum. (Sufficient Condition)
  - If $\forall\ x, H(x)$ is Positive Definite, then x^* is a strict global minimum. (Sufficient Condition)
  - If $\forall\ x, H(x)$ is Negative Semi-Definite, then x^* is a global maximum. (Sufficient Condition)
  - If $\forall\ x, H(x)$ is Positive Semi-Definite, then x^* is a global minimum. (Sufficient Condition)

3 Constrained Optimization I: Representing Constrained Sets

3.1 Some Examples and Issues

- max $f(x)$ over x s.t. $x \in C$.
- max x s.t. $x < 1$ has NO SOLUTION because the constraint set is NOT CLOSED.
- Sometimes you can substitute the constraint into the maximization and solve like an unconstrained problem. Even if you do find a solution like $f'(x^*) = 0$, remember ALWAYS check second order conditions (SOCs) for max or min.
- Sometimes there will be two solutions to a maximization, as is the case when the constraint set is not convex.

3.2 Functional Representations of Constraint Sets

- $f : \mathbb{R}^n \mapsto \mathbb{R}, C = \{x \in \Re^n : f(x) - c \geq 0\}$.

3.3 Open, Closed, and Bounded Sets

- **Definition:** A set of points $S \subset \Re^n$ is OPEN if $\forall\, x \in S$, $\exists$ an $\epsilon > 0$, s.t. $B_\epsilon(x) \subset S$. So (1,2) is open, but (1,2] is not open.

- **Theorem:** Any union of open sets is open.

- **Theorem:** Any finite intersection of open sets is open. Why finite? Consider $S_i = (-1/i, 1/i)$, $i = 1, 2, \dots$. Let $S = \cap_{i=1}^{n} S_i = (-1/n, 1/n)$. If $n \to \infty$, $\cap S_i = 0$. The singleton, zero, is NOT open.

- **Definition:** The INTERIOR of a set S is the largest open set contained in S. If S is $[1,2]$, $INT(S) = (1,2)$. Or $INT(S) = \cup_i S(i)$ where S_i is open and $S_i \subset S$.

- **Theorem:** If S is open, then $INT(S) = S$.

- **Definition:** Let $S \subset \Re^n$ and let $\{x_m\}$ be any sequence of elements in S. If $lim_{m\to\infty} x_m = r \in S\ \forall$ convergent sequences in S, then the set S is CLOSED.

- **Theorem:** A set S is CLOSED iff S^c is OPEN.

- Note that $\emptyset$ and $\Re_n$ are both OPEN and CLOSED by convention. The set $(1,2]$ is neither OPEN nor CLOSED. $[1,\infty)$ is CLOSED. Note we have to watch out when one side of the interval is infinite.

- **Definition:** The BOUNDRY of S, $Bd(S)$ is the set of all points such that $\forall\, \epsilon > 0$,

$$B_\epsilon(x) \cap S \neq \emptyset,$$

$$B_\epsilon(x) \cap S^c \neq \emptyset.$$

- Note that if S is OPEN, $Bd(S) \notin S$.

- **Theorem:** If S is closed, then $Bd(S) \subseteq S$.

- **Definition:** A set $S \in \Re^n$ is BOUNDED if $\exists\, M \in \Re_+$ such that $\forall\, x \in S, ||x|| \leq M$.

- **Definition:** A set $S \subset \Re^n$ is COMPACT if it is both CLOSED and BOUNDED.

- **Theorem:** (Weierstrass Theorem, SB 30.1) Let $f : \mathbb{R}^n \mapsto \mathbb{R}$ be continuous. If the set $C \subset \Re^n$ is closed and bounded (COMPACT), then $\exists\, x^*, y^* \in C$ such that $f(x^*) \geq f(x) \geq f(y^*)\ \forall\, x \in C$. In other words, x^* is a maximum and y^* is a minimum. Compactness gives us a lot. If C is compact, there will exists a solution to the problem $max_x f(x)$ s.t. $x \in C$. (And of course, the opposite minimization problem).

3.4 Convex Sets

- Define $C = \{x \in \Re^n : f(x) - c \geq 0\}$. If f is concave, then C is a convex set $\forall$ c. Same is true for several constraints which are all concave. The resulting constraint set (intersection) will be convex.

4 Constrained Optimization

4.1 Some Examples

- Typically we have $max_x f(x)$ s.t. $x \in C$ where,

$$C = \{x \in \Re^n : g_i(x_i) \geq 0 \text{ for } i = 1 \ldots m, g_i : \mathbb{R}^n \mapsto \mathbb{R}\}.$$

  OR,

$$C = \{x \in \Re^n : G(x) \geq 0, G : \mathbb{R}^n \mapsto \mathbb{R}^m\}.$$

  Where $G(x)$ is a vector valued function.

- NB: If constraint set is $G(x) \leq 0$, reverse it so $H(x) = -G(x) \geq 0$.

- NB: Non-Negativity constraints can be represented by: $g_i(x) = x_i \geq 0$.

- NB: Equality constraints can be represented by two constraints, one $\geq$ and one $\leq$.

A General Typology of Constrained Maximization Problems

- 1) Unconstrained Maximization: the constraint set, C, is just the whole space in which x lies ($\Re^n$ for example).

- 2) Lagrange Maximization Problems: C is only defined by equality constraints:

$$C = \{x \in \Re^n : G(x) = 0, G : \mathbb{R}^n \mapsto \mathbb{R}^m\}.$$

- 3) Linear Programming Problems. $g_i(x) = a_i'x - b_i$ $\forall$ $i = 1, 2, \ldots, m$ and $f(x) = c'x$. Everything is linear.

- Kuhn-Tucker Problems: Take a class of constraints (the non-negativity constraints) away from the constraint set C and describe these separately. So $max_x f(x)$ s.t. $G(x) \geq 0$, $x \geq 0$, $G : \mathbb{R}^n \mapsto \mathbb{R}^m$.

4.2 Lagrange Maximization Problems (LP)

- So here we $max_{x \in \Re^n} f(x)$ s.t. $G(x) = 0$, $G : \mathbb{R}^n \mapsto \mathbb{R}^m$.

- **Theorem:** (Lagrange Theorem). In the constrainted max problem, LP, suppose that f and G are C^1 and that the nxm matrix $\nabla G(x^*)$ has rank m. Then if x^* solves LP,

$\exists$ a vector, $\lambda^* \in \Re^m$ such that:

$$\nabla f(x^*) + \nabla G(x^*)\lambda^* = 0.$$

This only a NECESSARY condition. It can also be written:

$$\nabla f(x^*) + \sum_{i=1}^{m} \nabla g_i(x^*)\lambda_i^* = 0.$$

Because:

$$\nabla G(x^*)\lambda^* = (\nabla g_1, \nabla g_1, \ldots, \nabla g_m) \begin{bmatrix} \lambda_1^* \\ \lambda_2^* \\ \vdots \\ \lambda_m^* \end{bmatrix} = \sum_{i=1}^{m} \nabla g_i(x^*)\lambda_i^*.$$

- **Remark :** We can also write the first order necessary conditions (FONC) from the theorem as:

$$\nabla f(x^*) = -\sum_{i=1}^{m} \nabla g_i(x^*)\lambda_i^*.$$

  Which says that the gradient of the objective function at a solution can be written as a linear combination of the gradient of the constraints with weights, λ. So with one constraint, the gradient of the objective and the gradient of the constraint point in opposite directions, so $\nabla f(x^*) = -k\nabla G(x^*)$.

- **Remark:** No claims are made about the sign of λ^*, the lagrange multipliers. Could be positive or negative. In KT, this will change.

- **Remark:** The condition that $\nabla G(x^*)$ has rank m is a version of a constraint qualification (CQ). It means that $\nabla G(x^*)$ has m linearly independent columns, or all the constraints are independent.

- Another way to solve the problem is to define the Lagrangian:

$$\mathcal{L}(x, \lambda) = f(x) + \lambda' G(x).$$

  Where, $\mathcal{L} : \mathbb{R}^{n+m} \mapsto \mathbb{R}$. So this can be written:

$$\mathcal{L}(x, \lambda) = f(x) + \sum_{j=1}^{m} \lambda_j g_j(x).$$

  Note this is for MAXIMIZING with EQUALITY constraints. The signs change if we switch to minimizing and we have to do something different (ie, KT) if the constaints are not equality.

- The FONCs are now:

$$\frac{\partial \mathcal{L}(x^*, \lambda^*)}{\partial x_1^*} = 0,$$
$$\frac{\partial \mathcal{L}(x^*, \lambda^*)}{\partial x_2^*} = 0,$$
$$\vdots$$
$$\frac{\partial \mathcal{L}(x^*, \lambda^*)}{\partial x_n^*} = 0,$$

And

$$\frac{\partial \mathcal{L}(x^*, \lambda^*)}{\lambda_1^*} = 0,$$
$$\frac{\partial \mathcal{L}(x^*, \lambda^*)}{\lambda_2^*} = 0,$$
$$\vdots$$
$$\frac{\partial \mathcal{L}(x^*, \lambda^*)}{\lambda_m *} = 0.$$

Note the last set of conditions just give us our original constraints.

4.3 Kuhn-Tucker and Differentiability (KT)

- Maximization of the form $max_x f(x)$ s.t. $G(x) \geq 0$, $G : \mathbb{R}^n \mapsto \mathbb{R}^m$, $x \geq 0$.

- The Lagrangian in KT problems is now:

$$\mathcal{L}(x, \lambda) = f(x) + \lambda' G(x).$$

BUT, now, $\mathcal{L} : \mathbb{R}_+^{n+m} \mapsto \mathbb{R}$. Note the domain of $\mathcal{L}$ now is strictly NON-NEGATIVE.

- See graph in notes for an important plot of $f(x)$ which shows many local maxes and mins as well as possible maxes right on the borders of the domain space. You might get solutions to maximization problems where the derivative (gradient) of the function is not zero ... namely, right on the boundry of the constraint. $f(x) = x$ constrained on the set $[0, 10]$ achieves a maximum at $x^* = 10$, though $f'(x^*) = 1 \neq 0$. Thus KT theory allows us to check for both interior as well as boundry solutions.

- To set up the theory, we first need to define two sets of indices. Consider the KT problem with $f : \mathbb{R}^n \mapsto \mathbb{R}$ and $G : \mathbb{R}^n \mapsto \mathbb{R}^m$, and f and G are C^1. Then for any x^*, define the indices of G where the constraint binds and the indices of x where the non-negativity constraint DOES NOT bind. So:

$$K = \{i : g_i(x^*) = 0\},$$

$$M = \{i : x_i^* > 0\}.$$

So g_i binds but x_i is strictly greater than zero. Now construct the gradient of G by: 1) Only use binding constraints and 2) Only differentiate with respect to the x_i's that are strictly positive (DO NOT bind). Let $a_{ij} = \frac{\partial g_i}{\partial x_j}$ for $i \in K$ and $j \in M$. Denote the size (cardinality) of K as $|K|$ and $card(M) = |M|$. Then:

$$H(x^*) = \nabla_M G_K(x^*) = \begin{bmatrix} a_{11} & \dots & a_{|K|1} \\ \vdots & & \vdots \\ a_{1|M|} & \dots & a_{|K||M|} \end{bmatrix}.$$

- **Theorem:** (Kuhn Tucker (Karush)). Suppose x^* solves the Kuhn Tucker problem as a local max and suppose that $H(x^*)$ has maximum rank (Constraint Qualification). Then $\exists\ \lambda^* \in \Re^m_+$ such that:
  - a) $\frac{\partial \mathcal{L}(x^*, \lambda^*)}{\partial x_i} \leq 0$ for $i = 1...n$.
  - b) $\frac{\partial \mathcal{L}(x^*, \lambda^*)}{\partial x_i} * x_i = 0$ for $i = 1...n$.
  - c) $\frac{\partial \mathcal{L}(x^*, \lambda^*)}{\partial \lambda_j} = g_j(x^*) \geq 0$ for $j = 1...m$.
  - d) $\frac{\partial \mathcal{L}(x^*, \lambda^*)}{\partial \lambda_j} * \lambda_j = 0$ for $j = 1...m$.

  Conditions (b) and (d) are "Complimentary Slackness" conditions because :
  - If $\frac{\partial \mathcal{L}}{\partial x_i} < 0$ then $x_i^* = 0$ (So non-negativity constraint binds).
  - If $x_i^* > 0$ then $\frac{\partial \mathcal{L}}{\partial x_i} = 0$ (So non-negativity constraint does not bind, or it's SLACK).
  - If $g_j > 0$ then $\lambda_j^* = 0$ (the g_j constraint is slack).
  - If $\lambda_j^* > 0$ then $g_j = 0$ (the g_j constraint must bind).
- **Remark:** The condition that $H(x^*)$ has maximal rank is a version of the Constraint Qualification (CQ).
- **Remark:** Satisfying the KT FOCs is only a necessary condition for an optimum. We will still need more. SOCs.
- Often times, you solve the FONCs and then look at cases, such as, x strictly in the interior of the constraint set, x on the border, etc. See notes for example.
- **Theorem:** Suppose $G(x)$ is concave and $f(x)$ is strictly quasi-concave (or $G(x)$ is strictly concave and $f(x)$ is quasi-concave), then if x^* solves the KT problem, x^* is UNIQUE. If $C = \{x : G(x) \geq 0, x \geq 0\}$ is COMPACT (closed and bounded)

and NON-EMPTY, and the objective function is continuous, then THERE EXISTS x^* which solves the problem. ——— The second part of this comes directly from Weierstrass theorem, but the first part is proved in the notes. Note that at least one of the functions (f or G) must be strict.

- To check for minimums and maximums, construct the "special" bordered hessian as :

$$BH = \begin{bmatrix} 0 & \cdots & 0 & \frac{\partial g_1}{\partial x_1} & \cdots & \frac{\partial g_1}{\partial x_n} \\ \vdots & & \vdots & \vdots & & \vdots \\ 0 & \cdots & 0 & \frac{\partial g_k}{\partial x_1} & \cdots & \frac{\partial g_k}{\partial x_n} \\ \frac{\partial g_1}{\partial x_1} & \cdots & \frac{\partial g_k}{\partial x_1} & \frac{\partial^2 \mathcal{L}}{\partial x_1^2} & \cdots & \frac{\partial^2 \mathcal{L}}{\partial x_1 \partial x_n} \\ \vdots & & \vdots & \vdots & & \vdots \\ \frac{\partial g_1}{\partial x_n} & \cdots & \frac{\partial g_k}{\partial x_n} & \frac{\partial^2 \mathcal{L}}{\partial x_n \partial x_1} & \cdots & \frac{\partial^2 \mathcal{L}}{\partial x_n^2} \end{bmatrix}. \tag{3}$$

So we have zeros in the NW corner, the gradient of the binding constraints in the NE and SW corners, and the hessian of the Lagrangian in the SE corner. The whole damn thing is $(n+k)x(n+k)$. The condition is as follows: If the last $n-k$ leading principal minors alternate in sign with the sign of the LARGEST (Check on this!!!) equal to $(-1)^n$, then the BH is negative definite at x^* and x^* is a strict local max. If the last $n-k$ l.p.m.'s all have sign equal to $(-1)^k$, the BH is positive definite and x^* is a strict local min.

- Finally, see notes for an example where we find the solution (strict maximum) but we find that the CQ fails even when x^* is the solution. We could either ignore the Constraint Qualification, or look for points that violate CQ and include them as candidates for a solution.

5 Some Applications

- **Theorem:** Suppose x^* solves $max_{x \in \Re^n} f(x)$ s.t. $G(x) \geq 0$ and $x \geq 0$. Suppose $g : \mathbb{R} \mapsto \mathbb{R}$ is strictly monotonic. Further suppose $h : \mathbb{R}^n \mapsto \mathbb{R}$ s.t. $h(x) = g(f(x))$. THEN x^* solves $max_{x \in \Re^n} h(x)$ s.t. $G(x) \geq 0$ and $x \geq 0$.

- So any monotonic transformation of the objective function will result in the same solution to the maximization (minimization) problem.

5.1 Consumer Choice

- If the constraint sets involve a union of two sets, you may have to split the max problem into two sections over the different domain spaces.

5.2 Cost Minimization

- See notes for an example where the solution is that the Marginal Rate of Technical Solution (MRTS) is equal to the ratio of the input prices.

5.3 Exam Question 2000

- Recall that if the objective function is continuous and the contraint set is COMPACT (closed and bounded) and non-empty, Weiestrass says that a solution EXISTS. If in addition, the objective function is strictly concave and the constraint set is convex, then the solution is UNIQUE and thus the FONCs of the KT problem are SUFFICIENT.

6 Comparative Statics

6.1 The Implicit Function Theorem

- Suppose we have the following problem: $max_{x\in\Re^n} f(x)$ s.t. $G(x) \geq 0$, $x \geq 0$. Assume f, G are C^1 and assume the CQ holds. We would like to see how the optimum changes when we change a parameter. So rewrite, $max_{x\geq 0} f(x;q), G(x;q) \geq 0$.

- Denote the solution to this problem as $x(q)$. Define the value of the objective function at the optimum:

$$V(q) = f(x(q), q).$$

  This is the VALUE function of the problem. How does V changes with q ?

- **Definition:** The endogenous variable is an EXPLICIT function of the exogeneous variables:

$$y = F(x_1, x_2, \ldots, x_n).$$

- **Definition:** If for each $(x_1, x_2, \ldots, x_n)$, the following determines a corresponding value of y, then we say that the following defines y as an IMPLICIT function of $(x_1, x_2, \ldots, x_n)$:

$$G(x_1, x_2, \ldots, x_n, y) = 0.$$

- Suppose there is a C^1 solution $y = y(x)$ for $G(x, y) = 0$, where $G(x, y)$ is an implicit function of y. That is $G(x, y(x)) = 0$ solves. Totally differentiating with respect to x yields:

$$\frac{\partial G(x_0, y(x_0))}{\partial x} + \frac{\partial G(x_0, y(x_0))}{\partial y}\frac{\partial y(x_0)}{\partial x} = 0.$$

$$\frac{\partial G(x_0, y(x_0))}{\partial x} + \frac{\partial G(x_0, y(x_0))}{\partial y} y'(x_0) = 0.$$

  Thus:

$$y'(x_0) = -\frac{\partial G(x_0, y(x_0))/\partial x}{\partial G(x_0, y(x_0))/\partial y}.$$

So if the solution exists and is C^1, then it is necessary that:

$$\frac{\partial G(x_0, y(x_0))}{\partial y} \neq 0.$$

- **Theorem:** (SB 15.1) Let $G(x,y)$ be C^1 on a ball around $(x_0, y_0) \in \Re^2$. Suppose $G(x_0, y_0) = 0$. If:

$$\frac{\partial G(x_0, y(x_0))}{\partial y} \neq 0,$$

  then there exists a C^1 function $y = y(x)$ defined on an interval I about or around x_0 such that the following is true:

  - 1) $G(x, y(x)) = 0 \; \forall \; x \in I$.
  - 2) $y(x_0) = y_0$.
  - 3) $y'(x_0) = -\dfrac{\partial G(x_0, y(x_0))/\partial x}{\partial G(x_0, y(x_0))/\partial y}$.

- **Theorem:** The Implicit Function Theorem (IFT) with n variables (z) and p parameters (q). (SB 15.7)
  Suppose $T : \mathbb{R}^{n+p} \mapsto \mathbb{R}^n$ is C^1. Let $T(z^*, q^*) = 0$. If the nxn matrix formed by stacking the n derivatives (w.r.t. z) of $T_1, T_2, \ldots, T_n$ is invertible, that is, if the matrix has full rank or its determinant is nonzero, then there exists n C^1 functions each mapping $\mathbb{R}^p \mapsto \mathbb{R}^n$ such that:

  - 1) $z_1(q^*) = z_1^*, z_2(q^*) = z_2^*, \ldots, z_n(q^*) = z_n^*$.
  - 2) $T(z(q), q) = 0 \; \forall \; q \in B_\epsilon(q^*)$, for some $\epsilon > 0$.
  - 3) The following derivative must be invertible:

$$-[D_z T(z(q), q)]^{-1} D_q T(z(q), q).$$

  Note that:

$$D_z T(z(q), q) = \begin{bmatrix} \partial T_1/\partial z_1 & \ldots & \partial T_1/\partial z_n \\ \vdots & & \vdots \\ \partial T_n/\partial z_1 & \ldots & \partial T_n/\partial z_n \end{bmatrix}.$$

6.2 Theorem of the Maximum

- Consider the Lagrange problem (LP), $V(q) = max_x f(x,q)$, s.t. $G(x,q) = 0$, where x is $nx1$, G is $mx1$ and q is $px1$. Define, $T : \mathbb{R}^{n+m+p} \mapsto \mathbb{R}^{n+m}$ by:

$$T(x, \lambda; q) = \begin{bmatrix} \nabla_x f(x,q) + \lambda' \nabla_x G(x,q) \\ G(x,q) \end{bmatrix}.$$

  Thus the FONC's are $T(x^*, \lambda^*, q) = 0$.

- If the matrix of derivatives of T with respect to x and λ is invertible, then we can find C^1 functions $x^*(q)$, $\lambda^*(q)$ s.t. $T(x^*(q), \lambda^*(q), q) = 0$ for q in a neighborhood of some given q. Denote this matrix as the following:

$$\nabla T(x, \lambda; q) = \begin{bmatrix} \nabla_x^2 f(x,q) + \lambda' \nabla_x^2 G(x,q) & \nabla_x G(x,q) \\ (\nabla_x G(x,q))' & 0 \end{bmatrix}.$$

- **Theorem:** Theorem of the Maximum: Suppose LP satisfies the conditions of the IFT at $(x^*(q^*), q^*)$. If f is C^1 at $(x^*(q^*), q^*)$ then $V(q)$ is C^1 at q^*.

6.3 The Envelope Theorem

- Consider the Unconstrained Problem (UP): $max_x f(x; a)$, s.t. $f : \mathbb{R}^n \mapsto \mathbb{R}$ and f is C^1. Suppose $a \in \Re$.

- **Theorem:** (Envelope) (SB 19.4) Let $x^*(a)$ be a solution to the UP. Suppose $x^*(a)$ is a C^1 function of a. Then:

$$\frac{df(x^*(a), a)}{da} = \frac{\partial f(x^*(a), a)}{\partial a}.$$

  Note the total differential of f with respect to a should involve the partial of f wrt x times the derivatives of x wrt a plus the partial of f wrt a. This first term is 0 at an optimum, so the total differential is just the partial of f wrt a, ie, the direct effect of a on f. The indirect effects are 0.

- Now consider a KT type problem. $max_x f(x, q)$ s.t. $G(x, q) \geq 0$, $G : \mathbb{R}^n \mapsto \mathbb{R}^m$. Assume at the solution, the FOC hold with EQUALITY (important assumption).

- **Theorem:** (SB 19.5) Suppose KT satisfies the conditions of the IFT at $(x(q^*), q^*)$. If f is C^1 at $(x(q^*), q^*)$, then:

$$\begin{aligned} \nabla_q V(q^*) &= \nabla f(x(q^*), q^*) \text{ [By Definition]}. \\ &= \frac{\partial \mathcal{L}(x(q^*), \lambda(q^*), q^*)}{\partial q}. \\ &= \frac{\partial f(x(q^*), q^*)}{\partial q} + \lambda(q^*) \frac{\partial G(x(q^*), q^*)}{\partial q}. \end{aligned}$$

  See notes for proof. Note we assumed the constraint binded at the optimum. This may be relaxed though we have not shown it yet.

- Examples: One example shows that if the problem was such that increasing a parameter meant that the constraint was relaxed, we found that the change in the value function from raising the parameter was equal to λ, or the PRICE OF THE CONSTRAINT. When a constraint does not bind, $\lambda = 0$, so the prices of relaxing the constraint (further) is 0.

- Utility Maximization. $V(q) = V(p, I) = max_x U(x)$ s.t. $I - p'x \geq 0$, with $q = (p, I)$, the parameters. Thus,
$$\mathcal{L} = U(x) + \lambda(I - p'x).$$
Thus,
$$\frac{\partial V(p, I)}{\partial I} = \frac{\partial \mathcal{L}}{\partial I},$$
by the envelope theorem. And thus,
$$\frac{\partial V(p, I)}{\partial I} = \frac{\partial \mathcal{L}}{\partial I} = \lambda(p, I),$$
or the Marginal Utility of Income. Here V is sometimes called the indirect utility function. The level of utility which solves the consumers maximization problem.

- **Definition:** ROY'S IDENTITY.
$$\frac{\partial V(p, I)}{\partial p_i} = \frac{\partial \mathcal{L}}{\partial p_i} = -\lambda(p_i, I)x_i.$$
Thus,
$$x_i = \frac{\partial V(p, I)/\partial p_i}{\lambda(p, I)} = \frac{\partial V(p, I)/\partial p_i}{\partial V(p, I)/\partial I}.$$
And this is Roy. x_i is the demand for good i.

- Cost Minimization. $C(y, w) = min_x w'x$ s.t. $f(x) - y \geq 0$ with $q = (y, w)$ the parameters. Thus,
$$\mathcal{L} = -w'x + \lambda(f(x) - y).$$
And using the envelope theorem again,
$$\frac{\partial C(y, w)}{\partial w_i} = -\frac{\partial \mathcal{L}}{\partial w_i} = x_i(y, w).$$
And this is SHEPHARDS LEMMA. x_i is the demand for input i.

- **Definition:** A FIXED POINT of a function, $f : \mathbb{R}^n \mapsto \mathbb{R}^n$ is a point, x^*, such that $x^* = f(x^*)$. Note that not all functions have fixed points and some have more than one.

- **Theorem:** (Brouwer's Fixed Point Theorem) Suppose $f : \mathbb{R}^n \mapsto \mathbb{R}^n$ and for some convex and compact set, $C \subset \Re^n$, f maps C into itself (if $x \in C$, then $f(x) \in C$). If f is continuous, then it has a fixed point.

DYNAMIC SECTION

7 Differential Equations

7.1 Basic Concepts for Univariate Equations

- **Definition:** An ordinary differential equation (ODE) is an equation of the following form:
$$x^m = f[t, x(t), x^1(t), x^2(t), \ldots, x^{m-1}(t); \alpha].$$
Where:
$$x_t : \mathbb{R} \mapsto \mathbb{R},$$
$$x^1(t) = \frac{dx(t)}{dt},\ x^2(t) = \frac{d^2x(t)}{dt^2},\ \ldots,\ x^m(t) = \frac{d^m x(t)}{dt^m}, \quad (1)$$
and $\alpha \in \Re^p$ is a vector of parameters. Thus $f : \mathbb{R}^{m+1+p} \mapsto \mathbb{R}$.

- **NB:** The solution to (1) is a function $x(t)$ that together with its derivatives, satisfies (1) for a given α.

- **NB:** We call (1) an ORDINARY difeq because $x(t)$ is a function of only one variable, t. If x was a function of more than one variable, it would be a partial difeq.

- **NB:** A difeq is LINEAR if f is linear in $x(t)$ and its derivatives.

- **NB:** A difeq is AUTONOMOUS if t does not appear as a separate argument of f, but enters only thru x.

- **NB:** The ORDER of a difeq is the order of the highest derivative of x that appears in it. (1) is m^{th} order.

- **NB:** Any difeq can be reduced to a 1^{st} order system by introducing new variables. Namely an n^{th} order difeq will require $n-1$ new variables to reduce it to a 1^{st} order system.

- **Definition:** A PARTICULAR solution to (1) is a differentiable function $x(t)$ that satisfies (1) for some sub-interval, I_0, of the domain of definition of t, I. The set of all particular solutions is the GENERAL solution, $x^g(t)$.

- Often times the solution to a difeq is NOT unique. This problem can be overcome by augmenting the difeq with a BOUNDRY CONDITION of the form:
$$x(t_0) = x_0.$$

- **Definition:** A BOUNDRY VALUE problem is a problem that is defined by a difeq:
$$x^1(t) = f[t, x(t)],$$

and a boundry condition:

$$x(t_0) = x_0, \quad (x_0, t_0) \in XxI.$$

- **Theorem:** The Fundamental Existance / Uniqueness Theorem.
  Let f be C^1 in some neighborhood of (x_0, t_0). Then in some subinterval, I_0 of I containing t_0, there is a UNIQUE solution $x(t)$ to the boundry value problem defined above.

- **Definition:** Consider the following autonomous difeq:

$$x^1(t) = f[x(t)], \ \ f : X \subset \Re \mapsto \Re.$$

  A STEADY STATE is a point $\overline{x} \in X$, s.t. $f(\overline{x}) = 0$. Note that a steady state may not exist for a function, and some functions may have more than one steady state.

- **Definition:** Let $\overline{x}$ be a steady state of the autonomous difeq,

$$x^1(t) = f[x(t)], \quad f : X \subset \Re \mapsto \Re.$$

  Then we say $\overline{x}$ is STABLE if for any $\delta > 0$,

$$||x(t_0) - \overline{x}|| < \delta \Longrightarrow ||x(t) - \overline{x}|| < \delta \, \forall \, t \geq t_0.$$

  So any solution that gets sufficiently close to $\overline{x}$, stays within δ of $\overline{x}$ from then on.

- **Definition:** A steady state is ASYMPTOTICALLY STABLE if :
  - $\overline{x}$ is stable.
  - δ can be chosen $\ni$ any solution that satisfies $||x(t_0) - \overline{x}|| < \delta$ will also satisfy:

$$lim_{t \to \infty} x(t) = \overline{x}.$$

- A steady state can be both locally and/or globally stable.

- If for all x in some neighborhood of a steady state, $\overline{x}$,

$$x(t) < \overline{x} \Rightarrow x^1(t) > 0,$$

  and,

$$x(t) > \overline{x} \Rightarrow x^1(t) < 0,$$

  then $\overline{x}$ is locally asymptotically stable. If the signs are reversed, then $\overline{x}$ is unstable.

- **Definition:** A steady state, $\overline{x_i}$ is locally asymptotically stable if:

$$f'[x(t)]\bigg|_{x(t)=\overline{x_i}} < 0.$$

A steady state, $\overline{x_i}$ is unstable if:

$$f'[x(t)]\Big|_{x(t)=\overline{x_i}} > 0.$$

And if:

$$f'[x(t)]\Big|_{x(t)=\overline{x_i}} = 0,$$

then plot and see what the story is. A point can be semi-stable, or stable from the right or left but unstable on the opposite side.

- **Definition:** Let $\overline{x}$ be a steady state of :

$$x^1(t) = f[x(t)], \quad f : X \subset \Re \mapsto \Re, \quad f \in C^1.$$

We say that $\overline{x}$ is a hyperbolic equilibrium if:

$$f'[x(t)]\Big|_{x(t)=\overline{x}} \neq 0.$$

- **Theorem:** (Grobman-Hartman #1). If $\overline{x}$ is a hyperbolic equilibrium, then there exists a neighborhood, U, of $\overline{x} \ni$ the difeq above is topologically equivalent to the linear equation:

$$x^1(t) = f'[x(t)]\Big|_{x(t)=\overline{x}} * [x(t) - \overline{x}] \text{ in } U.$$

This is because what we just wrote is the 1^{st} order taylor series approximation of f around $\overline{x}$. Namely,

$$\underbrace{f[x(t)]}_{x^1(t)} = \underbrace{f(\overline{x})}_{0} + f'[x(t)]\Big|_{x(t)=\overline{x}} * [x(t) - \overline{x}] + \underbrace{R[x(t) - \overline{x}]}_{Remainder \approx 0}.$$

An Application of the Solow Growth Model

- Output, Y, is produced using capital (K) and labor (L). Denote:

$$Y(t) = F[K(t), L(t)].$$

Assume F is C^1, constant returns to scale (homogeneous of degree 1 or linearly homogeneous), and suppose F has positive and diminishing marginal products, ie:

$$\frac{\partial F}{\partial K} > 0, \quad and \quad \frac{\partial^2 F}{\partial K^2} < 0.$$

So F is strictly concave.

- Let $y = Y/L$, $k = K/L$. Thus,

$$\begin{aligned} Y(t) &= F[K(t), L(t)]. \\ &= L(t)F[K(t)/L(t), 1]. \\ &= L(t)F[k(t)]. \\ Y(t)/L(t) &= F[k(t)]. \\ y(t) &= f[k(t)]. \end{aligned}$$

- Suppose that f satisifies the following INADA conditions:
  - $f(0) = 0$.
  - $f'(0) = \infty$.
  - $lim_{k\to\infty} f'(k) = 0$.
- Assume a closed economy with a constant savings rate, s. s is the fraction of output which is saved.
- Assume the change in the capital stock comes from investment which is equal to the output of the economy times the savings rate:

$$K'(t) = \dot{K} = I(t) = sY = sF[K(t), L(t)], \quad s \in [0, 1].$$

- Assume that labor grows exogeneously at rate n. Thus,

$$L(t) = L(0)e^{nt} = L_0e^{nt}.$$

- Thus:

$$\begin{aligned} K(t) &= K(t)\frac{L(t)}{L(t)} \\ &= k(t)L(t) \\ &= k(t)L_0e^{nt} \\ K'(t) &= k'(t)L_0e^{nt} + nk(t)L_0e^{nt} \\ \frac{K'(t)}{L(t)} &= k'(t)\underbrace{\frac{L_0e^{nt}}{L(t)}}_{1} + nk(t)\underbrace{\frac{L_0e^{nt}}{L(t)}}_{1} \\ \frac{K'(t)}{L(t)} &= k'(t) + nk(t) \\ sf([k(t)] &= k'(t) + nk(t) \quad [From\ above] \end{aligned}$$

Rearranging:

$$k'(t) = sf[k(t)] - nk(t).$$

Which is the capital stock law of motion. A first order ordinary differential equation.

- Note the steady state at $\dot{k} = 0$ implies:

$$sf[k(t)] = nk(t).$$

At the steady state, of course k is constant, K is growing at rate n because $k = K/L$ is constant and L grows at n, then K must grow at n. What about Y? Note:

$$sf(\overline{k}) = n\overline{k} \Longrightarrow \frac{s}{n} = \frac{\overline{k}}{f(\overline{k})} = \frac{K/L}{Y/L} = \frac{\overline{K}}{\overline{\overline{Y}}}.$$

So Y is also growing at rate n.

- By looking at the graph or equation, it is clear that $\overline{k}$ is asymptotically stable.

Solving Autonomous Linear Difeq

- Suppose:

$$x^1(t) = ax(t) + b.$$

All solutions to this can be written:

$$\underbrace{x^g(t)}_{General\ Solution} = \underbrace{x^c(t)}_{Complimentary\ Function} + \underbrace{x^p(t)}_{Particular\ Solution}.$$

Where $x^c(t)$ is the general solution to the homogeneous equation that is associated with our difeq (excluding constants) and $x^p(t)$ is ANY particular solution to the full non-homogeneous equation. See notes for method to solve. $x^c(t)$ is found by separating the differential to achieve $x^c(t) = ce^{at}$ and $x^p(t)$ is found using the steady state condition, $x^1(t) = 0$, which implies $x^p(t) = -b/a$. Thus,

$$x^g(t) = ce^{at} - \frac{b}{a} \ \ if \ a \neq 0.$$

$$x^g(t) = bt + c \ \ if \ a = 0.$$

- We can solve for c using an initial condition. Suppose $x(0) = x_0$. The solution becomes:

$$x^g(t) = (x_0 + b/a)ce^{at} - \frac{b}{a} \ \ if \ a \neq 0.$$

$$x^g(t) = bt + x_0 \ \ if \ a = 0.$$

- Stability Properties of the solution depend on a. If $a > 0$, $e^{at} \to \infty$ as $t \to \infty$ and $x(t)$ explodes (unstable). If $a < 0$, $e^{at} \to 0$ as $t \to \infty$ and $x(t)$ is asymptotically stable.

7.2 Linear Systems of Differential Equations

- Now we write the difeq as :

$$x^1(t) = Ax(t) + b,$$

where A is an nxn matrix and b is an $nx1$ vector.

Uncoupled Systems (A Diagonal)

- A is of the form:

$$A = \begin{bmatrix} a_{11} & & 0 \\ & \ddots & \\ 0 & & a_{nn} \end{bmatrix}$$

- In this case, we can solve each equation separatly because the system is completely uncoupled. Thus, given an initial condition of $x_i(0) = x_{i0}$,

$$x_i^g(t) = (x_{i0} + b_i/a_{ii})e^{a_{ii}t} - b_i/a_{ii} \;\; for\ i = 1...n.$$

Diagonalizable Linear Systems

- In this case we can make A diagonal if it is not.
- **Theorem:** If the matrix A has n linearly independent eigenvectors, $v_1, v_2, \ldots, v_n$, then the matrix $V = [v_1, v_2, \ldots, v_n]$ is invertible and $V^{-1}AV$ is a diagonal matrix with the eigenvalues of A along the diagonal. Thus:

$$V^{-1}AV = \Lambda = \begin{bmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{bmatrix}$$

- **NB:** If all the eigenvalues of A are distinct, then the n eigenvectors of A are linearly independent and therefore A is diagonalizable (Sufficient Condition).
- Solving using a similar method to before, we get the general solution to the problem as :

$$x^g(t) = \sum_{i=1}^{n} c_i V_i e^{\lambda_i t} - A^{-1}b.$$

Assuming A is invertible.

- The constants $c_1 \ldots c_n$ can be found using a set of boundry conditions (n of them).
- Eigenvalues/Eigenvectors of A satisfy:

$$(A - \lambda I_n)V = 0.$$

- Eigenvalues are the roots of the following equation:

$$det(A - \lambda I_n) = 0.$$

- Now consider the stability of:

$$x^g(t) = \sum_{i=1}^{n} c_i V_i e^{\lambda_i t} - A^{-1} b.$$

If all eigenvalues are real, then we have two cases:

  - If $\lambda_i < 0 \; \forall \; i$, then $e^{\lambda_i t} \rightarrow 0$ as $t \rightarrow \infty$, so system is asymptotically stable.
  - If $\lambda_i > 0$ for some i, then $e^{\lambda_i t} \rightarrow \infty$ as $t \rightarrow \infty$, so $x_i(t)$ is unstable unless $c_i V_{ij} = 0$.

- If some eigenvalues are complex, we have to do more. First, a short review of complex numbers. A complex number has the form:

$$a + ib, \quad where \; a, b \in \Re, \quad i = \sqrt{-1}.$$

So a is the real part and b is the imaginary part.

- **Definition:** The CONJUGATE of a complex number $a + bi$ is $a - bi$.
- **Definition:** The MODULUS of a complex number, r, is $\sqrt{a^2 + b^2}$.
- **Definition:** Consider the angle, θ in the complex plane shown in the notes. Note that:

$$tan(\theta) = b/a.$$
$$cos(\theta) = a/r.$$
$$sin(\theta) = b/r.$$

Thus, $a = r \; cos(\theta)$ and $b = r \; sin(\theta)$. So rewrite the complex number as :

$$r \; cos(\theta) + i \; r \; sin(\theta).$$

$$r(cos(\theta) + i \; sin(\theta)).$$

Note that (by EULERS formula), $e^{i\theta} = cos(\theta) + i \; sin(\theta)$. Thus,

$$r(cos(\theta) + i \; sin(\theta)) = re^{i\theta}.$$

- If A is a matrix with real elements, then its complex eigenvalues and eigenvectors must come in conjugate pairs.

- So start with the general form of the solution:

$$x^g(t) = \sum_{i=1}^{n} c_i V_i e^{\lambda_i t} - A^{-1} b.$$

And assume that the first two eigenvalues, $\lambda_1 = \alpha + i\mu$ and $\lambda_2 = \alpha - i\mu$ are complex. Write the eigenvectors as :

$$V_1 = d + if.$$

$$V_2 = d - if.$$

For the first complex eigenvector:

$$V_1 e^{\lambda_1 t} = (d + if)e^{(\alpha+i\mu)t}.$$

$$V_1 e^{\lambda_1 t} = (d + if)e^{\alpha t}e^{i\mu t}.$$

Via Euler:

$$V_1 e^{\lambda_1 t} = (d + if)e^{\alpha t}(cos(\mu t) + i\ sin(\mu t)).$$

$$V_1 e^{\lambda_1 t} = e^{\alpha t}(d\ cos(\mu t) - f\ sin(\mu t)) + ie^{\alpha t}(f\ cos(\mu t) + d\ sin(\mu t)).$$

Let $u(t) = e^{\alpha t}(d\ cos(\mu t) - f\ sin(\mu t))$, and $v(t) = e^{\alpha t}(f\ cos(\mu t) + d\ sin(\mu t))$, so:

$$V_1 e^{\lambda_1 t} = u(t) + iv(t).$$

By symmetry:

$$V_2 e^{\lambda_2 t} = u(t) - iv(t).$$

So,

$$x^g(t) = c_1(u(t) + iv(t)) + c_2(u(t) - iv(t)) + \sum_{i=3}^{n} c_i V_i e^{\lambda_i t} - A^{-1}b.$$

Which is still complex. So choose $c_1 = k_1 + ik_2$ and $c_2 = k_1 - ik_2$. Thus,

$$x^g(t) = 2k_1 u(t) + 2k_2 v(t) + \sum_{i=3}^{n} c_i V_i e^{\lambda_i t} - A^{-1}b.$$

Now let $\tilde{c}_1 = 2k_1$ and $\tilde{c}_2 = 2k_2$. Thus,

$$x^g(t) = \tilde{c}_1 u(t) + \tilde{c}_2 v(t) + \sum_{i=3}^{n} c_i V_i e^{\lambda_i t} - A^{-1}b.$$

This is REAL.

- The stability of this equation depends on α. The sin and cos in the solution only introduces cycles into the solutions. $e^{\alpha t}$ determines the stability if any eigenvalues are complex. α is the real part of the complex number. Thus:

$$if\ \alpha > 0, e^{\alpha t} \longrightarrow \infty,\ as\ t \longrightarrow \infty.$$

$$if\ \alpha < 0, e^{\alpha t} \longrightarrow 0,\ as\ t \longrightarrow \infty.$$

MOST Common System of Differential Equations, $n = 2$

- Consider: $x^1(t) = Ax(t) + b$ and WLOG (See notes), let $b = 0$. Note that A is $2x2$. So,

$$x^1(t) = Ax(t).$$

- The eigenvalues of A solve:

$$|A - \lambda I_2| = 0.$$

Or,

$$\begin{vmatrix} a_{11} - \lambda & a_{12} \\ a_{21} & a_{22} - \lambda \end{vmatrix} = 0.$$

Thus,

$$(a_{11} - \lambda)(a_{22} - \lambda) - a_{12}a_{21} = 0.$$

$$\lambda^2 - \lambda(\underbrace{a_{11} + a_{22}}_{Trace(A)}) + \underbrace{a_{11}a_{22} - a_{12}a_{21}}_{Det(A)} = 0.$$

$$\lambda^2 - tr(A)\lambda + det(A) = 0.$$

So,

$$\lambda_{1,2} = \frac{tr(A) \pm \sqrt{tr(A)^2 - 4det(A)}}{2}.$$

- NB : $tr(A) = \lambda_1 + \lambda_2$.
- NB : $det(A) = \lambda_1\lambda_2$.
- Recall the simple ($n = 2$) differential equation:

$$x^1(t) = Ax(t),$$

where A is nxn. Recall the eigenvalues of A are as follow:

$$\lambda_{1,2} = \frac{tr(A) \pm \sqrt{tr(A)^2 - 4det(A)}}{2}.$$

- We can say something about the eigenvalues depending on the trace and determinant of A. Consider the cases:
  - Case 1: Nodes.
    * 1.a) If $|A| > 0$ and $[tr(A)^2 - 4|A|] > 0$, then λ_1 and λ_2 are real, distinct, and they have the same sign.
      · If $tr(A) < 0, \lambda_1$ and λ_2 are both negative and we have a stable Node.
      · If $tr(A) > 0, \lambda_1$ and λ_2 are both positive and we have an unstable Node.
    * 1.b) If $[tr(A)^2 - 4|A|] = 0$, then $\lambda_1 = \lambda_2 = \frac{tr(A)}{2}$, which are real and repeated.
      · If $tr(A) < 0, \lambda_1 = \lambda_2 < 0$, implies Stable Node.
      · If $tr(A) > 0, \lambda_1 = \lambda_2 > 0$, implies Unstable Node.

- Case 2: Saddle Points.
  If $|A| < 0$, λ_1 and λ_2 will be real since $[tr(A)^2 - 4|A|] > 0$, and they will also be distinct with opposite signs. See graph in notes for the phase diagram.
- Case 3: Spiral Points.
  If $[tr(A)^2 - 4|A|] < 0$, and $tr(A) \neq 0$, then λ_1 and λ_2 are complex conjugates. It can be shown that the spirals are clockwise if $a_{21} < 0$ and the spirals are counter-clockwise if $a_{21} > 0$. Note the $tr(A) = \lambda_1 + \lambda_2 = \alpha + i\mu + \alpha - i\mu = 2\alpha$.
  * If $tr(A) < 0$, the eigenvalues have a negative real part so the spiral converges.
  * If $tr(A) > 0$, the eigenvalues have a positive real part so the spiral diverges.
- Case 4: Centers.
  Here $tr(A) = 0$ and $|A| > 0$, so we get pure imaginary eigenvalues. Specifically:

$$\lambda_1 = \frac{\sqrt{-4|A|}}{2}, \quad \lambda_2 = -\frac{\sqrt{-4|A|}}{2}.$$

Non-Diagonalizable Systems - Not Covered

7.3 Elements of Non-Linear Systems

- Consider,

$$x^1(t) = F[x(t)], \quad F : X \subset \Re^n \mapsto \Re^n, \quad F \textit{ nonlinear.}$$

- The most important thing we will do with nonlinear systems is draw their phase diagrams. For $n = 2$, we have:

$$x_1^1(t) = f_1[x_1(t), x_2(t)],$$

$$x_2^1(t) = f_2[x_1(t), x_2(t)].$$

- There are 3 steps to drawing the phase diagrams.
  - Step 1: Set $x_1^1(t) = 0$ and $x_2^1(t) = 0$ to obtain the phase lines. This means:

$$x_1^1(t) = 0 \Longrightarrow f_1[x_1(t), x_2(t)] = 0,$$

$$x_2^1(t) = 0 \Longrightarrow f_2[x_1(t), x_2(t)] = 0.$$

    So, since we are in (x_1, x_2) space, we are really looking for the slope of the $x_2(t)$ function with respect to $x_1(t)$. By the Implicit Function Theorem:

$$\frac{dx_2(t)}{dx_1} = -\frac{\partial f_1/\partial x_1}{\partial f_1/\partial x_2}.$$

$$\frac{dx_2(t)}{dx_1} = -\frac{\partial f_2/\partial x_1}{\partial f_2/\partial x_2}.$$

    Note we use each f function for each phase line. Solve this to determine the sign of the slope and plot the phase line for $x_i^1(t) = 0$.

– Use the orginal equations to obtain the arrows of motion on either side of the phase lines. So what we need is:

$$\frac{\partial x_1^1(t)}{\partial x_1(t)} = \left.\frac{\partial f_1(\cdot)}{\partial x_1(t)}\right|_{x_1^1(t)=0} \quad OR \quad \left.\frac{\partial f_1(\cdot)}{\partial x_2(t)}\right|_{x_1^1(t)=0}.$$

Note we can choose which variable to take the partial with respect to depending on what is easier. HOWEVER, if one of the phase lines for $x_i^1(t) = 0$ is vertical or horizontal, then we may have no choice.

– Step 3: Combine the phase lines and directions of movement for the two variables. Stable and unstable arms depend on eigenvalues.

- Finally, if we have a nonlinear system, we can linearize it using a first order taylor series approximation. Consider,

$$x^1(t) = F[x(t)], \;\; F : X \subset \Re^n \mapsto \Re^n, \quad F \; nonlinear.$$

Then,

$$x^1(t) = F[x(t)] = F[\overline{x}] + DF[x(t)]\Big|_{x(t)=\overline{x}} [x(t) - \overline{x}] + R[x(t) - \overline{x}].$$

Noting that $F[\overline{x}] = 0$ by definition and the remainder terms are approximately zero, we have:

$$x^1(t) = F[x(t)] = DF[x(t)]\Big|_{x(t)=\overline{x}} [x(t) - \overline{x}].$$

Where:

$$DF[x(t)]\Big|_{x(t)=\overline{x}} = \begin{bmatrix} \frac{\partial f_1}{\partial x_1} & \frac{\partial f_1}{\partial x_2} & \cdots & \frac{\partial f_1}{\partial x_n} \\ \frac{\partial f_2}{\partial x_1} & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \vdots \\ \frac{\partial f_n}{\partial x_1} & \cdots & \cdots & \frac{\partial f_n}{\partial x_n} \end{bmatrix}.$$

- **Definition:** Let $\overline{x}$ be a steady state of the non-linear system,

$$x^1(t) = F[x(t)], \;\; F : X \subset \Re^n \mapsto \Re^n, \quad F \; nonlinear,$$

then we say that $\overline{x}$ is a hyperbolic equilibrium if the derivative of F, evaluted at $\overline{x}$,

$$DF[x(t)]\Big|_{x(t)=\overline{x}},$$

has no eigenvalues with a zero real part.

- **Theorem:** (Grobman-Hartman #2) If $\overline{x}$ is a hyperbolic equilibrium of the system,

$$x^1(t) = F[x(t)], \quad F : X \subset \Re^n \mapsto \Re^n, \quad F \; nonlinear,$$

and F is C^1, then $\exists$ a neighborhood, U, of $\overline{x}$, $\ni$ the system is topologically equivalent to the linear system:

$$x^1(t) = F[x(t)] = DF[x(t)]\Big|_{x(t)=\overline{x}} [x(t) - \overline{x}]$$

in U. This means it has the same qualitative properties as the original (nonlinear) system (stability, etc).

8 Difference Equations

8.1 Basic Concepts for Univariate Equations

- **Definition:** An ordinary difference equation is an equation of the form:

$$x_{t+m} = F[t, x_t, x_{t+1}, x_{t+2}, \ldots, x_{t+m-1}; \alpha].$$

Where $\alpha \in \Re^p$ is a vector of parameters. $F : \mathbb{R}^{m+1+p} \mapsto \mathbb{R}$. Since x is only a function of t, it's ordinary. It's linear if it's linear in all the x's. And it's autonomous if it does not contain t by itself, nonautonomous otherwise.

- The order of a difference equation is the difference between the highest and lowest subscripts that appear. Thus our difference equation is $(t + m) - t = m, m^{th}$ order.

- Note we can rewrite the above difference equation (by subtracting m from every subscript):

$$x_t = F[t, x_{t-1}, x_{t-2}, \ldots, x_{t-m}].$$

Assume that $m = 1$, then,

$$x_t = F[t, x_{t-1}].$$

In general the solution to a difference equation is not unique unless we introduce boundry conditions. Let $x_{t_0} = x_0$. We can derive the solution as follows:

$$x_{t_0+1} = F(t_0 + 1, x_{t_0}),$$

$$x_{t_0+2} = F(t_0 + 2, x_{t_0+1}),$$

$$\vdots$$

Steady States and Stability

- Consider the following difference equation:

$$x_t = f[x_{t-1}] \quad where\ f : X \subset \Re \mapsto \Re.$$

A steady state is a fixed point, $\overline{x}$, that is:

$$f(\overline{x}) = \overline{x}.$$

See graph in notes but it is easy to see that when the phase line is above the 45 degree line, $x_t = f(x_{t-1}) > x_{t-1}$ so x_t is increasing. The opposite is of course true. Then we see stability if the slope around the fixed point (steady state) has certain properties. Namely:

  - If $|f^1(\overline{x})| < 1$, then $\overline{x}$ is asymptotically stable.
  - If $|f^1(\overline{x})| > 1$, then $\overline{x}$ is unstable.
  - If $|f^1(\overline{x})| = 1$, then $\overline{x}$ could go either way.

- **Definition:** Let $\overline{x}$ be a steady state of the following diffence equation:

$$x_t = f[x_{t-1}] \quad where\ f : X \subset \Re \mapsto \Re.$$

Assume f is C^1. We say that $\overline{x}$ is a hyperbolic equilibrium if $|f'(\overline{x})| \neq 1$.

- **Theorem:** (Grobman-Hartman #3)
If $\overline{x}$ is a hyperbolic equilibrium of the difference equation,

$$x_t = f[x_{t-1}] \quad where\ f : X \subset \Re \mapsto \Re,$$

then there exists a neighborhood, U, of the steady state $\overline{x}$ such that the difference equation is topologically equivalent to the linear equation:

$$x_t = \overline{x} + f'(x_{t-1})\Big|_{x_{t-1}=\overline{x}} [x_{t-1} - \overline{x}]$$

In U.

Note that this last equation is just the first order taylor series approximation to our original difference equation:

$$x_t = f(x_{t-1}) = f(\overline{x}) + f'(x_{t-1})\Big|_{x_{t-1}=\overline{x}} [x_{t-1} - \overline{x}] + R.$$

$f(\overline{x}) = \overline{x}$ by definition and the remainder is approximately 0, so we just have:

$$x_t = f(x_{t-1}) = \overline{x} + f'(x_{t-1})\Big|_{x_{t-1}=\overline{x}} [x_{t-1} - \overline{x}].$$

Solving Autonomous Linear Difference Equations

- Suppose:

$$x_t = ax_{t-1} + b.$$

All solutions to this equation are of the form:

$$\underbrace{x_t^g}_{General} = \underbrace{x_t^c}_{Complementary} + \underbrace{x_t^p}_{Particular}.$$

- The complementary function is the solution to the homogeneous part of the equation, $x_t = ax_{t-1}$. Iterate by plugging in successive values of x_t to get:

$$x_t^c = a^t x_0 = ca^t.$$

- A particular soluton can be found using the steady state, so $x_t = x_{t-1}$. Thus,

$$x_t = ax_t + b.$$

$$x_t = \frac{b}{1-a}.$$

So,

$$x_t^p = \frac{b}{1-a} \ if \ a \neq 1..$$

- Adding the boundry condition: $x_0 = \tilde{x}$, yields:

$$x_0 = ca^0 + \frac{b}{1-a}.$$

$$c = \tilde{x} - \frac{b}{1-a}.$$

- So the solution to the boundry value problem is:

$$x_t = (\tilde{x} - \frac{b}{1-a})a^t + \frac{b}{1-a} \ \ if \ a \neq 1.$$

$$x_t = bt + \tilde{x} \ \ if \ a = 1.$$

- The stability properties of

$$x_t = ax_{t-1} + b.$$

all depend on the magnitude of a.

  - If $|a| < 1$, $a^t \to 0$ as $t \to \infty$, so x_t is asymptotically stable.
  - If $|a| > 1$, $a^t \to \infty$ as $t \to \infty$, so x_t is unstable unless $x_0 = \frac{b}{1-a}$, or we start at the steady state.

- However there is some more information we know based on the sign of a:
  - If $a > 0$, ca^t has the same sign for all t and the equation converges or diverges monotonically.
  - if $a < 0$, ca^t changes sign depending on t so the equation will jump from one side of the steady state to the other (either converging or diverging).

Solving NON-Autonomous Linear Difference Equations

- Suppose:

$$x_t = ax_{t-1} + b_t.$$

Where b_t is a function of t, and is often called the Forcing Term.

- Complementary Function (Same as autonomous):

$$x_t^c = ca^t.$$

- Particular Solution is found not using the steady state because b now depends on t. We can either solve via the backward or forward solution.
  - Backward solution (substitute directly for x_{t-1}):

$$x_t = a^s x_{t-s} + \sum_{j=0}^{s-1} a^j b_{t-j}.$$

Letting $s \longrightarrow \infty$,

$$x_t^B = \sum_{j=0}^{\infty} a^j b_{t-j}.$$

IF $|a| < 1$ and $|b_t| < B \; \forall \, t$. Yielding a general solution:

$$x_t^g = c_B a^t + \sum_{j=0}^{\infty} a^j b_{t-j}.$$

  - Forward solution (Rewrite as $x_{t-1} = f(x_t)$ and move forward one period to get $x_t = f(x_{t+1})$). Solving iteratively yields:

$$x_t = \frac{1}{a^s} x_{t+s} - \sum_{j=0}^{s-1} \frac{1}{a^{j+1}} b_{t+j+1}.$$

Letting $s \longrightarrow \infty$,

$$x_t^F = -\frac{1}{a} \sum_{j=0}^{\infty} \left(\frac{1}{a}\right)^j b_{t-j}.$$

IF $|a| > 1$ and $|b_t| < B \; \forall \, t$. Yielding a general solution:

$$x_t^g = c_F a^t - \frac{1}{a}\sum_{j=0}^{\infty}\left(\frac{1}{a}\right)^j b_{t-j}.$$

- So use the backward solution if $|a| < 1$ and use the forward solution if $|a| > 1$.

Applications

- Capital Accumulation: $K_t = (1-\delta)K_{t-1} + I_{t-1}$. Using backward iteration yields:

$$K_t = (1-\delta)^s K_{t-s} + \sum_{j=0}^{s-1}(1-\delta)^j I_{t-j-1}.$$

- Household's Intertemporal Budget Constraint: $a_t = (1+r)a_{t-1} + y_t - c_t$. Using foward iteration yields:

$$\sum_{j=0}^{\infty}\left(\frac{1}{1+r}\right)^j c_{t+j} = (1+r)a_{t-1} + \sum_{j=0}^{\infty}\left(\frac{1}{1+r}\right)^j y_{t+j}.$$

Which is the agent's lifetime budget constraint of intertemporal budget constraint.

8.2 Linear Systems of Difference Equations

- Consider the following system:

$$x_t = Ax_{t-1} + b.$$

Where A is nxn.

Uncoupled Systems: Diagonal A

- Solve each equation separately. Thus, given initial conditions $x_{i0} = \tilde{x}_i$ for $i = 1...n$,

$$x_{it} = (\tilde{x}_i - \frac{b_i}{1-a_{ii}})a_{ii}^t + \frac{b_i}{1-a_{ii}}, \; for \; i = 1 \ldots n.$$

Diagonalizable Linear Systems

- A is not diagonal here but, $V^{-1}AV$ is diagonal. Thus:

$$\Lambda = V^{-1}AV = \begin{bmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{bmatrix}.$$

Where V is a matrix with the eigenvectors of A for its columns.

- Complementary Function, x_t^c.

$$\begin{aligned} x_t &= Ax_{t-1}. \\ V^{-1}x_t &= V^{-1}Ax_{t-1}. \\ V^{-1}x_t &= V^{-1}AVV^{-1}x_{t-1}. \\ y_t &= \Lambda y_{t-1}. \end{aligned}$$

With $y_t = V^{-1}x_t$. Thus,

$$y_{it} = c_i\lambda_i^t.$$

$$x_t^c = Vy_i = \sum_{i=1}^{n} V_i c_i \lambda_i^t.$$

- Particular Solution, x_p^t, use the steady state.

$$x_t = Ax_t + b.$$

$$x_t^p = (I - A)^{-1}b.$$

- General Solution:

$$x_t^g = \sum_{i=1}^{n} V_i c_i \lambda_i^t + (I - A)^{-1}b.$$

- Stability of this general solution depends on the magnitude of the eigenvalues.
  - Assume all eigenvalues are real.
    * If $|\lambda_i| < 1 \; \forall \; i$, $\lambda_i^t \to 0$, as $t \to \infty$, so asymptotically stable.
    * If $|\lambda_i| > 1$ for some i, $\lambda_i^t \to \infty$, as $t \to \infty$, so x_{jt} explodes unless $c_i V_{ij} = 0$.
- Assume some eigenvalues are complex.
  - Here we need to transform the solution into a real valued solution. Thus assume, λ_1 and λ_2 are complex and $\lambda_3 \ldots \lambda_n$ are real. So:

$$\lambda_1 = \alpha + i\mu \text{ and } V_1 = d + if.$$

$$\lambda_2 = \alpha - i\mu \text{ and } V_2 = d - if.$$

Recall:

$$\lambda_1 = \alpha + i\mu = r(cos(\theta) + i\ sin(\theta)).$$

$$\lambda_2 = \alpha - i\mu = r(cos(\theta) - i\ sin(\theta)).$$

Where:

$$tan(\theta) = \mu/\alpha, \quad r = \sqrt{\alpha^2 + \mu^2}.$$

– And the general solution is:

$$x_t^g = \tilde{c}_1 u_t + \tilde{c}_2 w_t + \sum_{j=3}^{n} c_i V_i \lambda_i^t + (I - A)^{-1} b,$$

with:

$$u_t = r^t(d\ cos(\theta t) - f\ sin(\theta t).$$
$$w_t = r^t(f\ cos(\theta t) + d\ sin(\theta t).$$

– So r determines the stability of the system.

* If $r < 1$, $r^t \to 0$, as $t \to \infty$, so asymptotically stable.
* If $r > 1$, $r^t \to \infty$, as $t \to \infty$, so unstable.

Non-Diagonalizable Linear Systems - Not Covered

8.3 Elements of Non-Linear Systems

- Suppose:

$$x_t = F(x_{t-1}), \quad F : X \subset \mathbb{R}^n \mapsto \mathbb{R}^n, \quad F \in C^1, \quad nonlinear.$$

- We can linearize this equation using a first order taylor approximation around the steady state, $\overline{x}$:

$$x_t = F(x_{t-1}) = \underbrace{F(\overline{x})}_{\overline{x}} + DF(x_{t-1})\Bigg|_{x_{t-1}=\overline{x}} [x_{t-1} - \overline{x}] + \underbrace{R}_{\approx 0}.$$

$$x_t = F(x_{t-1}) = \overline{x} + DF(x_{t-1})\Bigg|_{x_{t-1}=\overline{x}} [x_{t-1} - \overline{x}].$$

- **Definition:** Let $\overline{x}$ be a steady state of:

$$x_t = F(x_{t-1}), \quad F : X \subset \mathbb{R}^n \mapsto \mathbb{R}^n, \quad F \in C^1, \quad nonlinear.$$

We say $\overline{x}$ is a hyperbolic equilibrium if:

$$DF(x_{t-1})\Bigg|_{x_{t-1}=\overline{x}}$$

has no eigenvalues with modulus ($r = \sqrt{\alpha^2 + \mu^2}$) equal to 1.

- **Theorem:** (Grobman-Hartman #4) If $\overline{x}$ is a hyperbolic equilibrium of the difference equation:

$$x_t = F(x_{t-1}), \quad F : X \subset \mathbb{R}^n \mapsto \mathbb{R}^n, \quad F \in C^1, \quad nonlinear.$$

Then there exists some neighborhood, U, around $\overline{x}$ such that the difference equation (nonlinear) is topologically equivalent to the linear system:

$$x_t = F(x_{t-1}) = \overline{x} + DF(x_{t-1})\Big|_{x_{t-1}=\overline{x}} [x_{t-1} - \overline{x}] \text{ in } U.$$

- Craziness.
- Consider the nonlinear system:

$$x_t = F(x_{t-1}), \quad F : X \subset \mathbb{R}^n \mapsto \mathbb{R}^n, \quad F \in C^1, \quad \textit{nonlinear.}$$

- To draw the phase diagram of this difference equation system, first draw the phase lines. For $n = 2$, we have:

$$x_{1,t+1} = f_1(x_{1t}, x_{2t}),$$
$$x_{2,t+1} = f_2(x_{1t}, x_{2t}).$$

Subtract x_{1t} from both sides of the first equation:

$$x_{1,t+1} - x_{1t} = f_1(x_{1t}, x_{2t}) - x_{1t}.$$

$$\Delta x_{1t} = f_1(x_{1t}, x_{2t}) - x_{1t}.$$

And similarly,

$$\Delta x_{2t} = f_2(x_{1t}, x_{2t}) - x_{2t}.$$

Now draw the phase lines by setting $\Delta x_{1t} = 0$. This implies,

$$f_1(x_{1t}, x_{2t}) - x_{1t} = 0.$$

And we need $\frac{dx_{2t}}{dx_{1t}}$. So by the IFT:

$$\frac{dx_{2t}}{dx_{1t}} = -\frac{\partial f_1/\partial x_{1t}}{\partial f_1/\partial x_{2t}}.$$

Repeat for x_{2t} and plot the phase lines. To determine the direction of motion, consider,

$$\frac{\partial \Delta x_{1t}}{\partial x_1} = \frac{\partial f_1(x_{1t}, x_{2t})}{\partial x_1} - 1,$$

or,

$$\frac{\partial \Delta x_{1t}}{\partial x_2} = \frac{\partial f_1(x_{1t}, x_{2t})}{\partial x_2}.$$

at some point on the phase line (usually the steady state). Note again we can use either differential depending on what is easier or necessary.

9 Discrete Time Intertemporal Optimization

9.1 Alternative Methods

- Suppose we have the following problem:

$$max_{\{c_t\}_{t=0}^T} \sum_{t=0}^{T} \beta^t u(c_t), \quad (1)$$

Subject to:

$$a_{t+1} = (1+r)(a_t + y_t - c_t), \ \ t = 0, 1, \ldots, T, \quad (2)$$

$$\beta = \frac{1}{1+\rho},$$

$$\rho > 0,$$

$$r > 0,$$

$$a_0 \ given.$$

- And note that a_t is the assets at the beginning of period t. y_t is labor income at the beginning of t. c_t is consumption expenditure at the beginning of t. ρ is the time discount rate. r is the interest rate. u is C^2, strictly increasing, and strictly concave. Also assume:

$$lim_{c_t \to 0} u'(c_t) = \infty.$$

This last assumption guarantees we will NOT have a boundry solution.

Method 1: Substitution

- Here we'll substitute the constraint (2) into objective function (1) directly. Thus:

$$max_{\{a_{t+1}\}_{t=0}^T} \sum_{t=0}^{T} \beta^t u\Big[-\frac{a_{t+1}}{1+r} + a_t + y_t\Big].$$

Note that a_t is given at time t so maximize over a_{t+1}.

- FOC:

$$\beta^t u' \underbrace{\Big[-\frac{a_{t+1}}{1+r} + a_t + y_t\Big]}_{c_t}\Big(\frac{-1}{1+r}\Big) + \beta^{t+1} u' \underbrace{\Big[-\frac{a_{t+2}}{1+r} + a_{t+1} + y_{t+1}\Big]}_{c_{t+1}} = 0.$$

$$\Longleftrightarrow$$

$$\beta^t u'[c_t]\Big(\frac{-1}{1+r}\Big) + \beta^{t+1} u'[c_{t+1}] = 0.$$

$$\beta^t u'(c_t) = \beta^{t+1}(1+r)u'(c_{t+1}).$$

$$u'(c_t) = \beta(1+r)u'(c_{t+1}). \quad (4)$$

- Recall the agent's lifetime budget constraint that we have derived previously by iteration:
$$\sum_{t=0}^{T}\left(\frac{1}{1+r}\right)^t c_t + \left(\frac{1}{1+r}\right)^{T+1} a_{T+1} = a_0 + \sum_{t=0}^{T}\left(\frac{1}{1+r}\right)^t y_t. \quad (5)$$

- The no-Ponzi game condition says that the agent cannot die with negative assets (debt). A better assumption (and one that will fall out later), is that $a_{T+1} = 0$, or the agent has exactly no assets on the 'deathday'. So, the lifetime budget constraint becomes:
$$\sum_{t=0}^{T}\left(\frac{1}{1+r}\right)^t c_t = a_0 + \sum_{t=0}^{T}\left(\frac{1}{1+r}\right)^t y_t.$$

- Now suppose:
$$u(c_t) = \frac{c_t^{1-\sigma}}{1-\sigma}, \quad \sigma > 0.$$
This is a constant elasticity of substitution (CES) utility function. Thus,
$$u'(c_t) = \frac{(1-\sigma)c_t^{-\sigma}}{1-\sigma} = c_t^{-\sigma}.$$

- So (4) implies:
$$c_t^{-\sigma} = \beta(1+r)c_{t+1}^{-\sigma}.$$
$$c_{t+1}^{\sigma} = \beta(1+r)c_t^{\sigma}.$$
$$c_{t+1} = [\beta(1+r)]^{1/\sigma} c_t. \quad (6)$$
Assume:
$$\beta^{1/\sigma}(1+r)^{1/\sigma} < 1+r.$$
Divide by $1+r$,
$$\beta^{1/\sigma}(1+r)^{1/\sigma-1} < 1.$$

- Now (6) impiles:
$$c_t = [\beta(1+r)]^{t/\sigma} c_0. \quad (7)$$
Because if $x_t = ax_{t-1}$, the solution to this difference equation is $x_t = x_0 a^t$.

- Substituting (7), the solution to the difference equation on consumption, into (5), the lifetime budget constraint, yields:
$$\sum_{t=0}^{T}\left(\frac{1}{1+r}\right)^t \underbrace{[\beta(1+r)]^{t/\sigma} c_0}_{c_t} = a_0 + \sum_{t=0}^{T}\left(\frac{1}{1+r}\right)^t y_t.$$
$$c_0 \sum_{t=0}^{T}\left(\beta^{1/\sigma}(1+r)^{1/\sigma-1}\right)^t = a_0 + \sum_{t=0}^{T}\left(\frac{1}{1+r}\right)^t y_t.$$

Note that $\sum_{t=0}^{T} x^t = \dfrac{1-x^{T+1}}{1-x}$,

$$c_0 \frac{1-\left(\beta^{1/\sigma}(1+r)^{1/\sigma-1}\right)^{T+1}}{1-\left(\beta^{1/\sigma}(1+r)^{1/\sigma-1}\right)} = a_0 + \sum_{t=0}^{T}\left(\frac{1}{1+r}\right)^t y_t.$$

Solve for c_0:

$$c_0 = \frac{1-\left(\beta^{1/\sigma}(1+r)^{1/\sigma-1}\right)}{1-\left(\beta^{1/\sigma}(1+r)^{1/\sigma-1}\right)^{T+1}}\left[a_0 + \sum_{t=0}^{T}\left(\frac{1}{1+r}\right)^t y_t\right]. \quad (8)$$

- So (8) is our formula for consumption at time 0 and we can use (6) to find consumption in future periods.
- Finally, in the limit as $T \to \infty$,

$$c_0 = 1-\left(\beta^{1/\sigma}(1+r)^{1/\sigma-1}\right)\left[a_0 + \sum_{t=0}^{\infty}\left(\frac{1}{1+r}\right)^t y_t\right]. \quad (9)$$

Method 2: Lagrange

- Consider again the problem in (1) and (2) taking into account the terminal non-indebtedness condition (No Ponzi Games).

$$max_{\{c_t\}_{t=0}^T} \sum_{t=0}^{T} \beta^t u(c_t), \quad (18)$$

Subject To:

$$\sum_{t=0}^{T}\left(\frac{1}{1+r}\right)^t c_t + \left(\frac{1}{1+r}\right)^{T+1} a_{T+1} = a_0 + \sum_{t=0}^{T}\left(\frac{1}{1+r}\right)^t y_t. \quad (19)$$

$$\left(\frac{1}{1+r}\right)^{T+1} a_{T+1} \geq 0 \quad (20).$$

- We can combine (19) and (20) to get:

$$\underbrace{\sum_{t=0}^{T}\left(\frac{1}{1+r}\right)^t y_t - \sum_{t=0}^{T}\left(\frac{1}{1+r}\right)^t c_t}_{(\frac{1}{1+r})^{T+1} a_{T+1}} \geq 0. \quad (21)$$

- The lagrangian:

$$L = \sum_{t=0}^{T} \beta^t u(c_t) + \lambda\Big[\sum_{t=0}^{T}\Big(\frac{1}{1+r}\Big)^t y_t - \sum_{t=0}^{T}\Big(\frac{1}{1+r}\Big)^t c_t\Big]. \quad (22)$$

- FOCs:

$$\frac{\partial L}{\partial c_t} = \beta^t u'(c_t) - \lambda\Big(\frac{1}{1+r}\Big)^t = 0.$$

$$\Leftrightarrow \beta^t u'(c_t) = \lambda\Big(\frac{1}{1+r}\Big)^t. \quad (23)$$

$$\frac{\partial L}{\partial c_{t+1}} \Rightarrow \beta^{t+1} u'(c_{t+1}) = \lambda\Big(\frac{1}{1+r}\Big)^{t+1}. \quad (24)$$

$$\frac{\partial L}{\partial \lambda} = \sum_{t=0}^{T}\Big(\frac{1}{1+r}\Big)^t y_t - \sum_{t=0}^{T}\Big(\frac{1}{1+r}\Big)^t c_t \geq 0 \quad (25).$$

With Complimentary Slackness:

$$\lambda\frac{\partial L}{\partial \lambda} = \lambda\Big[\sum_{t=0}^{T}\Big(\frac{1}{1+r}\Big)^t y_t - \sum_{t=0}^{T}\Big(\frac{1}{1+r}\Big)^t c_t\Big] = 0. \quad (26)$$

- (24) / (23) implies:

$$u'(c_t) = \beta(1+r)u'(c_{t+1}) \;\; t = 0, 1, \ldots, T-1 \quad (27).$$

Notice that since (27) = (4), we have the same solution.

- Finally, note that if $t = 0$, this means that (23) becomes:

$$\lambda = u'(c_0) > 0,$$

by assumption. Thus if λ is positive, the constraint binds, which means,

$$\sum_{t=0}^{T}\Big(\frac{1}{1+r}\Big)^t y_t - \sum_{t=0}^{T}\Big(\frac{1}{1+r}\Big)^t c_t = \frac{1}{1+r}^{T+1} a_{T+1} = 0,$$

Or the No-Ponzi condition holds with equality. No assets when you die.

9.2 The Maximum Principal

- Define:
x_t : A stock variable (state variable) which is measured at the beginning of period t.
u_t : A flow variable (control variable).
$x_{t+1} - x_t = f_t(x_t, u_t)$ ... the evolution of the stock variable.
$g_t(x_t, u_t) \geq 0$ ... other inequality constraints.

- Assume the objective function is additively separable:

$$\sum_{t=0}^{T} r_t(x_t, u_t).$$

All variables are taken at time t.

- The optimization problem is going to be:

$$max_{\{u_t\}_{t=0}^{T} \ \{x_t\}_{t=1}^{T}} \quad \sum_{t=0}^{T} r_t(x_t, u_t), \quad (1)$$

subject to:

$$\lambda_t : x_{t+1} - x_t = f_t(x_t, u_t), \ \ t = 0, 1, \ldots, T \quad (2),$$

$$\mu_t : g_t(x_t, u_t) \geq 0, \ \ t = 0, 1, \ldots, T \quad (3),$$

$$x_0 \text{ is given.} \quad (4)$$

$$x_{T+1} \text{ is given.} \quad (5)$$

- Write the lagrangian:

$$L = \sum_{t=0}^{T} r_t(x_t, u_t) + \lambda_t[f_t(x_t, u_t) + x_t - x_{t-1}] + \mu_t g_t(x_t, u_t), \quad (6)$$

with $\lambda_t \in \Re, \mu_t \in \Re_+$.

- FOCs:

$$\frac{\partial L}{\partial u_t} = \frac{\partial r_t}{\partial u_t} + \lambda_t \frac{\partial f_t}{\partial u_t} + \mu_t \frac{\partial g_t}{\partial u_t} = 0, \ \ t = 0, ..., T. \quad (7)$$

$$\frac{\partial L}{\partial x_t} = \frac{\partial r_t}{\partial x_t} + \lambda_t \frac{\partial f_t}{\partial x_t} + \lambda_t - \lambda_{t-1} + \mu_t \frac{\partial g_t}{\partial x_t} = 0, \ \ t = 1, ..., T. \quad (8)$$

$$\frac{\partial L}{\partial \lambda_t} = f_t(u_t, x_t) + x_t - x_{t-1} = 0 \ \ \forall\, t \quad (9).$$

$$\frac{\partial L}{\partial \mu_t} = g_t(u_t, x_t) \geq 0, \ \ \mu_t \frac{\partial L}{\partial \mu_t} = \mu_t g_t(u_t, x_t) = 0 \ \ \forall\, t \quad (10).$$

- Define a function H_t, called the Hamiltonian:

$$H_t(x_t, u_t, \lambda_t) = \underbrace{r_t(x_t, u_t)}_{Instantaneous\ Return} + \lambda_t \underbrace{f_t(x_t, u_t)}_{Evolution\ of\ Stock}.$$

- Note that (7) says that u_t should be chosen to maximize H_t s.t. $g_t(x_t, u_t) \geq 0$. Denote $H_t^*(x_t, u_t, \lambda_t)$ the resulting maximum value.

- The lagrangian for this single-period optimization probem is:

$$\mathcal{L} = H_t(x_t, u_t, \lambda_t) + \mu_t g_t(x_t, u_t). \quad (12)$$

Using (12), rewrite (8) as:

$$\frac{\partial \mathcal{L}}{\partial x_t} + \lambda_t - \lambda_{t-1} = 0.$$

$$-\frac{\partial \mathcal{L}}{\partial x_t} = \lambda_t - \lambda_{t-1}.$$

By the envelope theorem:

$$\lambda_t - \lambda_{t-1} = -\frac{\partial H_t^*}{\partial x_t}. \quad (13).$$

Also by the envelope theorem:

$$\frac{\partial H_t^*}{\partial \lambda_t} = \frac{\partial \mathcal{L}}{\partial \lambda_t} = f_t.$$

- (9) can be rewritten:

$$\underbrace{\frac{\partial H_t^*}{\partial \lambda_t}}_{f_t} + x_t - x_{t+1} = 0.$$

$$x_{t+1} - x_t = \frac{\partial H_t^*}{\partial \lambda_t}.$$

- So finally, we get to the MAXIMUM PRINCIPAL. The first order necessary conditions for the problem in (1) to (5) are :
  - i) For each t, u_t maximizes the Hamiltonian, $H_t(u_t, x_t, \lambda_t)$ such that the single period constraints $g(x_t, u_t) \geq 0$. This is like:

$$\frac{\partial \mathcal{L}}{\partial u_t} = 0.$$

  - ii) For each t,

$$x_{t+1} - x_t = \frac{\partial H_t^*}{\partial \lambda_t}.$$

  - iii) For each t,

$$\lambda_t - \lambda_{t-1} = -\frac{\partial H_t^*}{\partial x_t}.$$

- **Remark 1:** If x_{T+1} is not given and there is only a non-negativity restriction on the terminal value of the state variable, $x_{T+1} \geq 0$, then the FONC's need to be augmented by the "Transversality Condition" which says:

$$\lambda_T \; x_{T+1} = 0. \quad (15)$$

That is, if you leave stock at the end, the value, λ_T (the shadow price) must be zero, or the stock must be worthless.

- **Remark 2:** For infinite horizon problems, the FONCs remain unchanged, but (15) should be replaced by :

$$lim_{T\to\infty}\lambda_T \;\; x_{T+1} = 0.$$

- **Remark 3:** The FONC's are sufficient for a unique optimum, if the appropriate conditions are imposed on our functions, r, f and G. (ie, concave functions maximized over convex and compact constraint sets).

- Finally, consider equation (8):

$$\underbrace{[\frac{\partial r_t}{\partial x_t} + \mu_t \frac{\partial g_t}{\partial x_t}]}_{Marginal\ Effect\ of\ x\ this\ period} \quad + \quad \underbrace{\lambda_t \frac{\partial f_t}{\partial x_t}}_{Effect\ next\ period} \quad + \underbrace{\lambda_t - \lambda_{t-1}}_{Capital\ Gain} = 0.$$

So the first two terms are like the stock's dividend and the last is the capital gain. Thus one of our conditions is that there is no excess return from holding a stock x – NO ARBITRAGE CONDITION.

9.3 Dynamic Programming

- Consider again the optimization problem:

$$max_{\{u_t\}_{t=0}^T} \;\; \sum_{t=0}^{T} r_t(x_t, u_t), \quad (1)$$

subject to:

$$x_{t+1} - x_t = f_t(x_t, u_t), \;\; t = 0, 1, \ldots, T \quad (2),$$

$$g_t(x_t, u_t) \geq 0, \;\; t = 0, 1, \ldots, T \quad (3),$$

$$x_0 \text{ is given.} \quad (4)$$

$$x_{T+1} \text{ is given.} \quad (5)$$

- **Definition:** Define the VALUE function as the resulting maximum value of the objective function expressed as a function of the inital state variables:

$$V_0(x_0).$$

- Now consider we are at time $t = \tau$. Then the value function, $V_\tau(x_\tau)$, is the value function of:

$$max_{\{u_t\}_{t=\tau}^T} \;\; \sum_{t=\tau}^{T} r_t(x_t, u_t),$$

subject to:

$$x_{t+1} - x_t = f_t(x_t, u_t), \;\; t = \tau, \tau+1, \ldots, T$$

$$g_t(x_t, u_t) \geq 0, \quad t = \tau, \tau+1, \ldots, T.$$

- Thus we have **Bellman's Principal.** Consider the choice u_t. This determines x_{t+1}. Then it remains to solve the subproblem that starts at period $t+1$ and achieves the maximum value $V_{t+1}(x_{t+1})$. So,

$$V_t(x_t) = max_{u_t}\left\{ r_t(x_t, u_t) + V_{t+1}(x_{t+1}) \right\}.$$

  This is Bellman's Equation.
  The constraints are now:

$$x_{t+1} - x_t = f_t(x_t, u_t), \quad \text{for a given } t$$

$$g_t(x_t, u_t) \geq 0, \quad \text{for a given } t$$

$$x_0 \text{ is given.}$$

$$x_{T+1} \text{ is given.}$$

- So in period T:

$$V_T(x_T) = max_{u_T}\left\{ r_T(x_T, u_T) + \underbrace{V_{T+1}(x_{T+1})}_{0} \right\}.$$

$$V_T(x_T) = max_{u_T}\left\{ r_T(x_T, u_T) \right\}.$$

  Subject to:

$$x_{T+1} - x_T = f_T(x_T, u_T),$$

$$g_T(x_T, u_T) \geq 0,$$

$$x_T,\ x_{T+1} \text{ given.}$$

  Solving this ($FOC = 0$) yields a POLICY function $h_T(x_T) = u_T^*$. Substituting this back into the value function gives us $V_T(x_T)$.

- And in period $T-1$:

$$V_{T-1}(x_{T-1}) = max_{u_{T-1}}\left\{ r_{T-1}(x_{T-1}, u_{T-1}) + V_T(x_T) \right\}.$$

  Subject to:

$$x_T - x_{T-1} = f_{T-1}(x_{T-1}, u_{T-1}),$$

$$g_{T-1}(x_{T-1}, u_{T-1}) \geq 0,$$

$$x_{T-1},\ x_{T+1} \text{ given.}$$

Note that $V_T(x_T)$ is from the previous iteration. Solving this $(FOC = 0)$ yields a POLICY function $h_{T-1}(x_{T-1}) = u^*_{T-1}$. Substituting this back into the value function gives us $V_{T-1}(x_{T-1})$.

- Continue this backwards until you reach period $t = 0$. In order for this to work, we need the instantaneous return function and the constraint functions to be such that controls at time t influence states, x_{t+s}, and returns, r_{t+s}, for $s > 0$ but not earlier. This is called TIME CONSISTANCY.

- Note that the Bellman and the Maximum Principal should both yield the same result!!

- See notes for an example. We start at time T, find the value function and the policy functions and then repeat once or twice. You start to see a pattern and can quickly write down a formula for $V_0(x_0)$.

- The next logical extension to the analysis is to ask what happens when T goes to ∞. Clearly, we cannot solve backwards because there is no end. Note that in the finite version, the value function was different at each time period, hence the t subscript on V. In that case, there was a different number of periods remaining for each time interval. When T goes to ∞, the value function is TIME INVARIANT, ie, it looks the same at all time periods. So as $T \to \infty$,

$$V_t(x_t) = V(x_t) \; \forall \; t.$$

This is all true ONLY under certain conditions to be discussed later.

Solving Intertemporal Optimization Problems with an Infinite Horizon

Method 1: Guess and Verify

- Proceed as follows:
  - 1) Guess a form for the value function. For instance, we could believe that the value function is of the form:

    $$V^G(x) = A + Bh(x) + Cz(x).$$

    For known functions $h(x), z(x)$, but unknown coefficients, A, B, and C.
  - 2) Plug the conjecture, $V^G(x)$ into BOTH sides of the Bellman equation:

    $$V^G(x_t) = max_{u_t}\left\{ r_t(x_t, u_t) + V^G(\underbrace{f(x_t, u_t) + x_t}_{x_{t+1}}) \right\}.$$

  - 3) Obtain the policy function from the FOC and plug it back into the bellman's equation.
  - 4) Find values of the coefficients (A, B, C) that make the equation hold.

- In practice, this guess will be given to us and we will need to verify that it works. We can also take the limit of the finite solution to find the infinite horizon solution.

Method 2: Successive Approximation

- Denote $V^j(x)$ the j^{th} guess of the value function, $V(x)$. Proceed as follows:
  - 1) Start with an initial guess $V^0(x)$, which is arbitrary... we can choose anything.
  - 2) Plug $V^0(x)$ into the RIGHT HAND SIDE ONLY of the Bellman's equation to generate a new function on the left hand side:

$$V^1(x_t) = max_{u_t}\left\{ r_t(x_t, u_t) + V^0(\underbrace{f(x_t, u_t) + x_t}_{x_{t+1}}) \right\}.$$

  - 3) Obtain the policy function from the FOC and plug it back into the above equation.
  - 4) If $V^1(x) = V^0(x)$, then the guess is correct and $V(x) = V^0(x)$.
  - 5) If $V^1(x) \neq V^0(x)$, then the guess was wrong and repeat. Plug $V^1(x)$ into the RHS to get a $V^2(x)$ on the LHS.

$$V^2(x_t) = max_{u_t}\left\{ r_t(x_t, u_t) + V^1(\underbrace{f(x_t, u_t) + x_t}_{x_{t+1}}) \right\}.$$

  - 6) Repeat until things converge ... under what conditions will a unique solution be found? More soon.

- So under what conditions is the value function time invariant and the sequence of functions $V^j(x)$ converge to a unique solution, $V^\infty(x)$, that solves the Bellman's equation?

- Consider again,

$$max_{\{u_t\}_{t=0}^T} \sum_{t=0}^{T} r_t(x_t, u_t), \quad (1)$$

  subject to:

$$x_{t+1} - x_t = f_t(x_t, u_t), \;\; t = 0, 1, \ldots, T \quad (2),$$
$$g_t(x_t, u_t) \geq 0, \;\; t = 0, 1, \ldots, T \quad (3),$$
$$x_0 \text{ is given.} \quad (4)$$
$$x_{T+1} \text{ is given.} \quad (5)$$

- For the optimization problem, let $T \to \infty$ and assume:
  - $r_t(x_t, u_t) = \beta^t r(x_t, u_t), \beta \in (0, 1)$. – Time Invariant.
  - $f_t(x_t, u_t) = f(x_t, u_t)$. – Time Invariant.

- $g_t(x_t, u_t) = g(x_t, u_t)$. – Time Invariant.

- Then the current value function is a function of the initial state only: $V(x_t)$, and we have:

$$V(x_t) = max_{u_t}\left\{ r(x_t, u_t) + \beta V(x_{t+1}) \right\}. \quad (14)$$

Subject to:

$$f(x_t, u_t) = x_{t+1} - x_t, \quad (15)$$
$$g(x_t, u_t) \geq 0. \quad (16)$$

- Also, the policy function, $u_t^* = h(x_t)$ is time invariant. Solving (14-16) is equivalent to finding a fixed point, V, of an operator, K, such that $V = KV$ where:

$$KV(x_t) = max_{u_t}\left\{ r(x_t, u_t) + \beta V(x_{t+1}) \right\}. \quad (17)$$

Subject to (15-16). We have to define a bunch of stuff to make this make sense. Basically under a few assumptions, if we can show that the operator, K, has certain properties, then we'll converge to a unique solution. But first some definitions and theorems.

- **Definition:** A METRIC space is a set, X, and a function, d, called a metric, $d : XxX \mapsto \Re$ such that $\forall\ (x, y, z) \in X$:
  - a) $d(x, y) \geq 0$ and $d(x, y) = 0$ iff $x = y$ – Positivity.
  - b) $d(x, y) = d(y, x)$ – Symmetry.
  - c) $d(x, y) \geq d(x, z) + d(z, y)$ – Triangle Inequality.

- **Definition:** A sequence $\{x_n\}$ in a metric space (X, d) is said to CONVERGE to a limit $x_0 \in X$ if:

$$\forall\ \epsilon > 0, \exists\ N(\epsilon) \ni \quad d(x_n, x_0) < \epsilon\ \forall\ n \geq N(\epsilon).$$

- **Definition:** A sequence $\{x_n\}$ in a metric space (X, d) is said to be CAUCHY CONVERGENT (or just a cauchy sequence) if:

$$\forall\ \epsilon > 0, \exists\ N(\epsilon) \ni \quad d(x_n, x_m) < \epsilon\ \forall\ n, m \geq N(\epsilon).$$

- **Theorem:** Every convergent sequence in a metric space is also a cauchy sequence. See notes for proof.

- The converse is not necessarily true. This leads to another definition. **Definition:** A metrix space (X, d) is COMPLETE is every cauchy sequence in the metric space converges to some point in X.

- **Definition:** Define the "SUP NORM" metric space as the set of continuous bounded functions mapping the interval $[0,1]$ into $\Re$ with:

$$d_\infty(x,y) = sup|x(j) - y(j)|.$$

Note the sup=supremum is the lowest upper bound.

- **Definition:** A function, $K : X \mapsto X$, that maps a metric space into itself is called an OPERATOR.

- **Definition:** Let (X,d) be a metric space and $K : X \mapsto X$ an operator in the metric space. K is a CONTRACTION if $\exists\ \beta \in [0,1)$ such that:

$$d(K(x), K(y)) \geq \beta d(x,y) \ \ \forall (x,y) \in X.$$

So basically K brings points closer together. β is called the modulus of K.

- **Theorem:** The Contraction Mapping Theorem. Let (X,d) be a COMPLETE metric space and $K : X \mapsto X$, a contraction in the space. Then:
  - i) There is a unique point, $x^* \in X$, such that $K(x^*) = x^*$ (a fixed point).
  - ii) The sequence $\{x_n\}$ defined by $x_1 = K(x_0), x_2 = K(x_1), \ldots, x_{n+1} = K(x_n)$ converges to x^* (for any starting point, x_0.)

- So returning to our initial problem:

$$KV(x_t) = max_{u_t}\Bigg\{ r(x_t,u_t) + \beta V(x_{t+1}) \Bigg\}. \quad (17)$$

Subject to (15-16). By the Contraction Mapping theorem, we just need to show that K is a contraction and we're golden. Because if we start at V^0 (our initial guess), then $V^1 = KV^0, V^2 = KV^1, \ldots, V^* = KV^*$.

- **Theorem:** Blackwell's Sufficient Conditions for a Contraction. Let K be an operator defined on a metric space (X, d_∞) where X is a space of functions. If K satisfies:
  - 1) Montonicity: For any $x,y \in X$, if $x \geq y$, then $K(x) \geq K(y)$.
  - 2) Discounting: $\exists\ \beta \in [0,1) \ni$

$$K(x+c) \leq K(x) + \beta c \ \forall\ x \in X, c \in \Re_+.$$

Then K is a CONTRACTION with modulus β. See notes for proof.

- In our problem: Assume $r(x_t,u_t)$ is real valued, continuous, concave, and bounded, and the constraint set:

$$\{x_t, x_{t+1}, u_t : x_{t+1} = f(x_t,u_t) + x_t, g(x_t,u_t) \geq 0\},$$

is convex and compact. We work with the metric space of continuous, bounded, real valued functions. The metric is given by d_∞. This metric space can be shown to be COMPLETE and also K (defined by 17) maps a continuous bounded function V into a continuous bounded function KV. So all we really need now is monotonicity and discounting.

- Does K satisfy Blackwell's conditions? Consider $V(x) \geq W(x)$ for all $x \in X$. Then,

$$\begin{aligned} KV(x_t) &= max_{u_t}\left\{ r(x_t, u_t) + \beta V(x_{t+1}) \right\} \;\; s.t.(15-16) \\ &\geq max_{u_t}\left\{ r(x_t, u_t) + \beta W(x_{t+1}) \right\} \;\; s.t.(15-16) \\ &= KW(x). \\ V(x) \geq W(x) &\Rightarrow KV(x) \geq KW(x). \end{aligned}$$

  So K is MONOTONIC.

- For any $c \in \Re_+$,

$$\begin{aligned} K(V(x_t)+c) &= max_{u_t}\left\{ r(x_t, u_t) + \beta V(x_{t+1}+c) \right\} \;\; s.t.(15-16) \\ &= max_{u_t}\left\{ r(x_t, u_t) + \beta V(x_{t+1}) + \beta c \right\} \;\; s.t.(15-16) \\ &= \underbrace{max_{u_t}\left\{ r(x_t, u_t) + \beta V(x_{t+1}) \right\}}_{KV(x)} + \beta c \;\; s.t.(15-16) \\ &= KV(x_t) + \beta c \\ K(V(x_t)+c) &= KV(x_t) + \beta c \end{aligned}$$

  SO K DISCOUNTS.

- Thus, K is a CONTRACTION, and we have all we need to guarantee that we will converge to a unque solution.

10 Continuous Time Intertemporal Optimization

10.1 The Maximum Principal

- Consider:

$$max_{\{u_t\}_{t=0}^T} \quad \sum_{t=0}^{T} r_t(x_t, u_t), \quad (1)$$

subject to:

$$x_{t+1} - x_t = f_t(x_t, u_t), \;\; t = 0, 1, \ldots, T \quad (2),$$

$$g_t(x_t, u_t) \geq 0, \;\; t = 0, 1, \ldots, T \quad (3),$$

$$x_0 \text{ is given.} \quad (4)$$

$$x_{T+1} \text{ is given.} \quad (5)$$

- Now let $\Delta t \to 0$. (2) becomes:

$$x(t + \Delta t) - x(t) = f(x(t), u(t), t)\Delta t.$$

Divide both sides by Δt and let $\Delta t \to 0$,

$$x^1(t) = f(x(t), u(t), t). \quad (7)$$

(3) becomes:

$$g(x(t), u(t), t) \geq 0. \quad (8)$$

(1) becomes:

$$\sum_{i=0}^{T/\Delta t} r(x(i\Delta t), u(i\Delta t), i\Delta t)\Delta t.$$

Let $\Delta t \to 0$,

$$\int_0^T r(x(t), u(t), t)dt. \quad (9)$$

- So the continuous time analog to (1-5) is:

$$max \int_0^T r(x(t), u(t), t)dt \quad (10),$$

Subject to:

$$x^1(t) = f(x(t), u(t), t), \quad (11)$$

$$g(x(t), u(t), t) \geq 0, \quad (12)$$

$$x(0) \; given \quad (13)$$

$$x(T) \; given \quad (14)$$

- Recall the the FOCs from the discrete time case:

$$\frac{\partial L}{\partial u_t} = \frac{\partial r_t}{\partial u_t} + \lambda_t \frac{\partial f_t}{\partial u_t} + \mu_t \frac{\partial g_t}{\partial u_t} = 0, \ \ t = 0, ..., T.$$

$$\frac{\partial L}{\partial x_t} = \frac{\partial r_t}{\partial x_t} + \lambda_t \frac{\partial f_t}{\partial x_t} + \lambda_t - \lambda_{t-1} + \mu_t \frac{\partial g_t}{\partial x_t} = 0, \ \ t = 1, ..., T.$$

$$\frac{\partial L}{\partial \lambda_t} = f_t(u_t, x_t) + x_t - x_{t-1} = 0 \ \ \forall\, t$$

$$\frac{\partial L}{\partial \mu_t} = g_t(u_t, x_t) \geq 0, \ \ \mu_t \frac{\partial L}{\partial \mu_t} = \mu_t g_t(u_t, x_t) = 0 \ \ \forall\, t$$

Their continuous time analog:

$$\frac{\partial L}{\partial u(t)} = \frac{\partial r(t)}{\partial u(t)} + \lambda(t) \frac{\partial f(t)}{\partial u(t)} + \mu(t) \frac{\partial g(t)}{\partial u(t)} = 0, \quad (15)$$

$$\frac{\partial L}{\partial x(t)} \Rightarrow \lambda^1(t) = -\left[\frac{\partial r(t)}{\partial x(t)} + \lambda(t) \frac{\partial f(t)}{\partial x(t)} + \mu(t) \frac{\partial g(t)}{\partial x(t)} \right] \quad (16)$$

$$\frac{\partial L}{\partial \lambda(t)} \Rightarrow x^1(t) = f(x(t), u(t), t) \quad (17)$$

$$\frac{\partial L}{\partial \mu(t)} = g(u(t), x(t), t) \geq 0, \ \ \mu(t) \frac{\partial L}{\partial \mu(t)} = \mu(t) g(u(t), x(t), t) = 0. \quad (18)$$

- Define the Hamiltonian in continuous time:

$$H(x(t), u(t), \lambda(t), t) = r(x(t), u(t), t) + \lambda(t) f(x(t), u(t), t). \quad (19)$$

- FOCs:

$$\frac{\partial H(t)}{\partial u(t)} + \mu(t) \frac{\partial g(t)}{\partial u(t)}. \quad (20)$$

$$\lambda^1(t) = -\left[\frac{\partial H(t)}{\partial x(t)} + \mu(t) \frac{\partial g(t)}{\partial x(t)} \right]. \quad (21)$$

$$x^1(t) = \frac{\partial H(t)}{\partial \lambda(t)} = f(x(t), u(t), t). \quad (22)$$

$$g(u(t), x(t), t) \geq 0, \ \ \mu(t) g(u(t), x(t), t) = 0. \quad (23)$$

- The MAXIMUM PRINCIPAL states that the FONCs of the problem in (10-14) are given by (20-23).

- **Remark 1:** If $x(T+1)$ is not given and there is only a non-negativity restriction on the terminal value of the state variable, $x(T+1) \geq 0$, then the FONC's need to be augmented by the "Transversality Condition" which says:

$$\lambda(T) \;\; x(T+1) = 0. \quad (15)$$

That is, if you leave stock at the end, the value, $\lambda(T)$ (the shadow price) must be zero, or the stock must be worthless.

- **Remark 2:** For infinite horizon problems, the FONCs remain unchanged, but (15) should be replaced by :
$$lim_{T\to\infty}\lambda(T) \;\; x(T+1) = 0.$$

- **Remark 3:** The FONC's are sufficient for a unique optimum, if the appropriate conditions are imposed on our functions, r, f and G. (ie, concave functions maximized over convex and compact constraint sets).

- Finally, we can distinguish between two types of hamiltonians (present value and current value). If the instantaneous return function is of the form:
$$r(x(t), u(t), t) = e^{-\rho t} r(x(t), u(t)),$$
the present value hamiltonian should be used which is of the form:
$$H(x(t), u(t), \lambda(t), t) = e^{-\rho t} r(x(t), u(t)) + \lambda(t) f(x(t), u(t), t).$$
So the exponential term is our discount factor in continuous time and notice that the return function is now time invariant.

- The first two FOCs (replacing 20-21) become:
$$e^{-\rho t}\frac{\partial r(t)}{\partial u(t)} + \lambda(t)\frac{\partial f(t)}{\partial u(t)} + \mu(t)\frac{\partial g(t)}{\partial u(t)}. \quad (24)$$
$$\lambda^1(t) = -\left[e^{-\rho t}\frac{\partial r(t)}{\partial x(t)} + \lambda(t)\frac{\partial f(t)}{\partial x(t)} + \mu(t)\frac{\partial g(t)}{\partial x(t)}\right]. \quad (25)$$
So $\lambda(t)$ is the value of the stock discounted back to period 0.

- The CURRENT value Hamiltonian is as follows:
$$\tilde{H}(x(t), u(t), \eta(t), t) = e^{\rho t} \underbrace{H(x(t), u(t), \lambda(t), t)}_{Present\ Value\ Hamiltonian}.$$
$$\tilde{H}(x(t), u(t), \eta(t), t) = r(x(t), u(t)) + e^{\rho t}\lambda(t) f(x(t), u(t), t).$$
$$\tilde{H}(x(t), u(t), \eta(t), t) = r(x(t), u(t)) + \eta(t) f(x(t), u(t), t).$$
So $\eta(t) = e^{\rho t}\lambda(t)$ is the value of the stock discounted back to period t.

- The first two FOCs (replacing 20-21) become:

$$\frac{\partial \tilde{H}(t)}{\partial u(t)} + \gamma(t)\frac{\partial g(t)}{\partial u(t)} = 0, \quad (27)$$

$$\eta^1(t) = \rho\eta(t) - \left[\frac{\partial \tilde{H}(t)}{\partial x(t)} + \gamma(t)\frac{\partial g(t)}{\partial x(t)}\right] \quad (28).$$

With $\gamma(t) = e^{\rho t}\mu(t)$.

11 Applications

11.1 Political Business Cycles

- Problem:

$$max \int_0^T v(u,p)e^{rt}dt,$$

subject to:

$$p = \phi(u) + a\pi,$$

$$\dot{\pi} = b(p - \pi),$$

$$\pi(0) \; given.$$

- Here the electorate votes according to the vote function, v, which depends on the unemployment rate, u, and the inflation rate, p. We also have an expectations augmented phillips curve (the first constraint), and adaptive inflation expectations, the second constraint. Note the discount term is actually positive so the politician values periods closer to the election more than today. Assume the following functional forms:

$$v(u,p) = -u^2 - hp,$$

$$p = (j - ku) + a\pi.$$

We can substitute these functions into the maximization problem and we just have the adaptive expectations contstraint. The state variable is π (we'll always be given the evolution of the state variable), and the control variable is u.

- Thus, the present value Hamiltonian is:

$$H = (-u^2 - hj + hku - ha\pi)e^{rt} + \lambda[b(j - ku - (1-a)\pi)].$$

- FOCs:

$$\frac{\partial H}{\partial u} = 0.$$

$$\dot{\lambda} = -\frac{\partial H}{\partial \pi}.$$

$$\dot{\pi} = \frac{\partial H}{\partial \lambda}.$$

Transversality Condition (TVC):

$$\lambda(T)\pi(T) = 0.$$

- Solving yields an equation for $u^*(t)$. We see in the notes that u^* is decreasing in t which makes sense because since the politicians value the time periods closer to the next election more than those close to $t = 0$ (voters are myopic), u^* should start out high and then fall throughout the term until the election.

11.2 Utility Maximization

- Problem:

$$max \sum_{t=0}^{\infty} \beta^t u(c_t),$$

subject to:

$$a_{t+1} = (1 + r)(a_t + y - c_t),$$

$$a_0 \; given.$$

- We know that the value function will be time invariant because the objective function is time invariant, the limits are infinite, and the constraint functions are time invariant. The state variable here is a_t. The choice can be either c_t or a_{t+1}. The Bellman is therefore:

$$V(a_t) = max_{c_t}\left\{ u(c_t) + \beta V(a_{t+1}) \right\},$$

s.t.

$$a_{t+1} = (1 + r)(a_t + y - c_t),$$

$$a_0 \; given.$$

Choose a_{t+1} as our choice variable. The Bellman becomes:

$$V(a_t) = max_{a_{t+1}}\left\{ u(\underbrace{-\frac{a_{t+1}}{1+r} + a_t + y}_{c_t}) + \beta V(a_{t+1}) \right\}.$$

- Take the FOC and solve for $u'(c_t)$. We will have a term like $V^1(a_{t+1})$. To find this, first evaluate $V^1(a_t)$ and then update. Note only take the derivative wrt to a_t directly. We can do this via the envelope theorem. The indirect effects are all 0. Solving this will yield $u(c_t)$ in terms of $u(c_{t-1})$ which is an EULER EQUATION.

- Note the same result can be achieved if we used c_t as our choice variable. The EULER equation is:

$$u'(c_t) = \beta(1 + r)u'(c_{t+1}).$$

The marginal utility of consumption today must be equal to not consuming it today and waiting until tomorrow (with interest and discounting).

- So it appears that if have a continuous time problem (maximize the objective function over the integral from $t = 0$ to $t = \infty$), then use the hamiltonian - theorem of the maximum. If we have an infinite horizon discrete time problem, use dynamic programming (bellman equation).
- Finally, the Cauchy Schwarz Inequality:

$$\left(\sum_{k=1}^{n} a_k b_k \right)^2 \leq \sum_{k=1}^{n} a_k^2 \sum_{k=1}^{n} b_k^2.$$