ES202 Lecture Notes #1

1. Distribute course outline and signup list; office hours and TF
   75 minutes or 50+25 with break
   Book and tapes
   starting on the hour
   homework rules, quizzes, project, final
   recitation time for the T.F.

2. Difficulties of the course
   •M21ab & Stat 110 (needed after 11/15) or equivalent
   •linear independence of LODE solution, matrix of rank n has n linearly independent vectors. Use of probability
   •UG have successfully taken it in the past
   •cannot be crammed in the last minute
   •grading difficulty=0 , 1/3 A's, 2/3 B+ and B's no C's

3. Two principal questions of science and technology: Descriptive vs Prescriptive.
   Frame of mind for rational decision making

4. Examples of Optimization and Optimal Control

   (Cliff & Vincent JOTA 11/73)

$dy/dt=\alpha y(y^e-y)$	Verhust model of animal population	
y=fish population	y^e=equilibrium population	
nonlinear ODE:	for $y<<y^e$	$dy/dt=(\alpha y^e)y$ ==>exponential growth
	for y>>ye	$dy/dt=-\alpha y^2$ ==> $y(t) =1/\alpha t$ rapid decay

   controlled growth
   $dy/dt=\alpha y(y^e-y) -vy$ where v=effort at harvesting fish $v\in[0,1]$

   criterion of performance: determine v(t) for t_o t t_f to
   Max $J=\phi(y(t_f))+ (c_1vy-c_2y-c_3v)dt$

   CURRENT MASS. GEORGE’S BANK FISH DEPLETION PROBLEM

   (i) dynamics — short vs long term considerations. Today's action effects the future
   (ii) performance criterion
   (iii) existence of control variable — emphasis on solution not on models
   (iv) qualitative vs. quantitative answers

5. Generalization to more complex food chain cycle and dynamics
   •multiple species of fish $y=[y_1, . . . , y_n]^T$

   $$dy_i/dt =f_i(y_1, . . . , y_n, t) \text{ or }$$
   $$=f_i(. . . y_j(t-\tau) . . .) \quad i=1,2, . . . , n$$

   •uncertainty since y must be estimated from observations, i.e.,
   z(t)=y(t)+ noise or more generally

=h(y(t), noise(t))
dy/dt=f(y) + distrubances

6. Optimal Resource Allocation example (Burghes Lyle Nicholos OCAM 3, 282-291, 1982)
dK/dt = f(K, R)-C where K=capital stock, R =resource usage, C=consumption
dS/dt=-R S(t) 0 S=nonrenewable resources

Initial state Ko and So given. Terminal condition $S(t_f)=0$ $K(t_f)=K_f$ = amount of new techn ology capital stock. Find R(t) and C(t) to maximize J
Max J= $e^{-\delta t}U(C)dt$
typical $U=(C-C_K)^{\beta}$ $f=K^{\alpha}R^{1-\alpha}$ $0<\alpha,\beta<1$

7. Other example: Apollo moon landing, desert storm war, peak load pricing, rent control, dam control, etc.

8. CHARACTERISTICS

Prescriptive	vs	Descriptive
Dyanmic	vs	Static
Quantitative	vs	Qualitative
Centralized	vs	Decentralized
Single criterion	vs	Multi-objective
Physical	vs	man-made
continuous	vs	combinatorial / discrete
concept`	vs	rigor

9. Notations (not theory) about linear algebra (see appendix):

10. Fundamental of static continuous optimization; continuity==> approximations are possible

(i) linear approximation
(ii) quaddratic approximations and higher via Taylor series expansion

$f(u) = f(u_o) + f/ u (u-u_o) +1/2 (u-u_o)^T\ ^2f/ x\ ^2(u-u_o)$ + higher order terms
all partials evaluated at u=uo

note for multivariate optimization, control variables interact, e.g., economic growth as function of discount rate and tax rate.

11. Concept of a gradient: dJ=???du

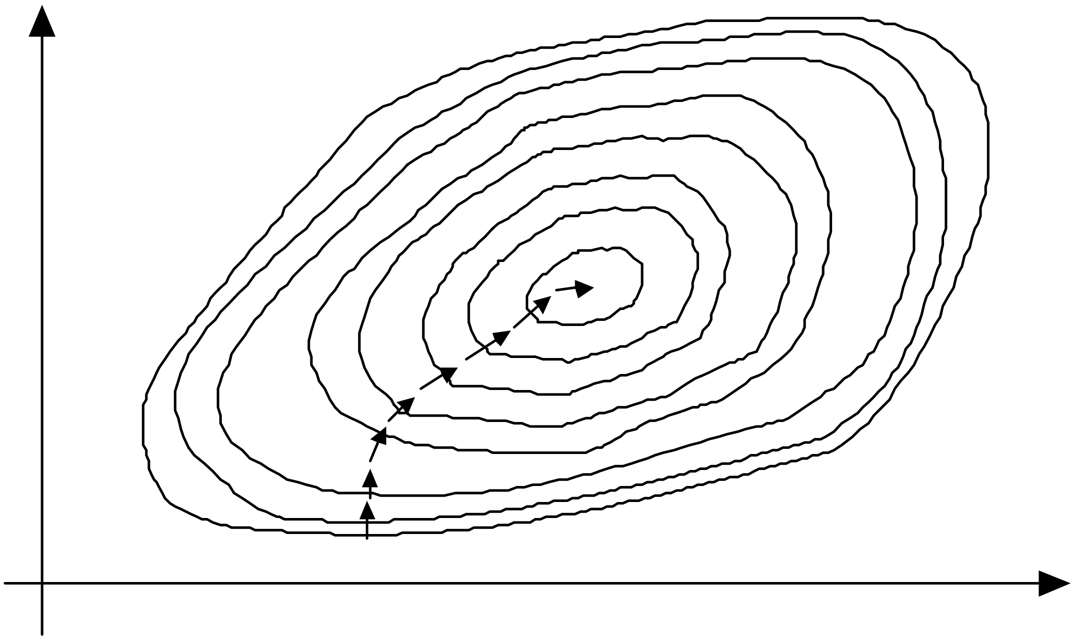
to first order ???= J/ u = [J/ u_1, ..., J/ u_m]

optimality <==> dJ 0 for all ***admissible*** du

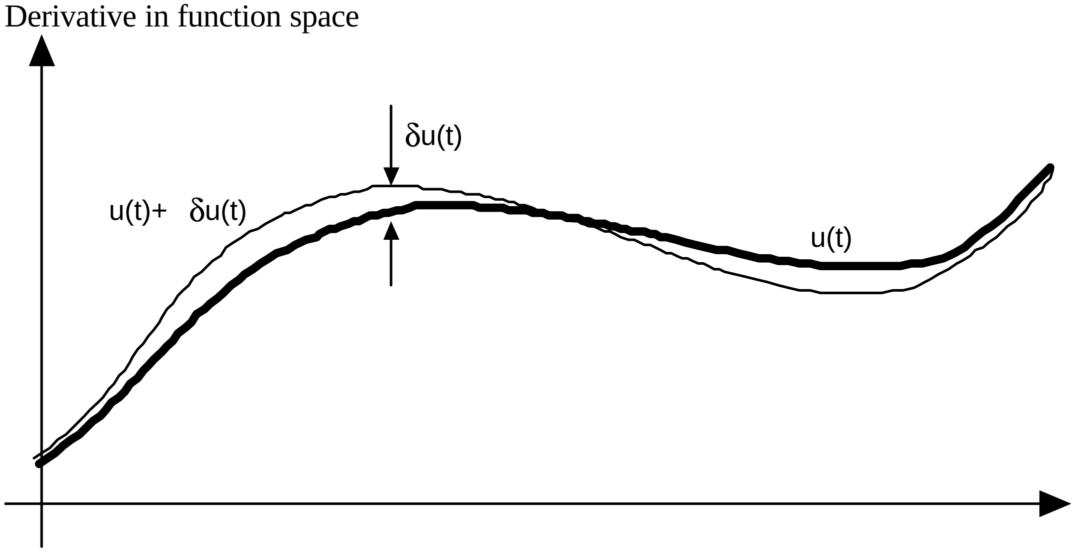
where admissible is taken to mean that du satisfies whatever constraint they may be.

Necessary condition $J/\ u = 0$
Sufficient condition $du^T[\ {}^2J/\ u\ {}^2]du>0$

12. Concept of successive approximation or skiing downhill in a thick fog



13. Derivative in function space



14. Peak load Electricity Pricing Model - to even out energy usage since electric energy cannot be easily stored

$\dot{q}$ = energy using rate $\quad$ =u (consumption rate)

$$J=\frac{1}{2}(q(T)-\bar{q})^2+\frac{a}{2}\int_0^T [(u-\bar{u})^2+p(t)u]dt \quad ; u(t) \bullet 0$$

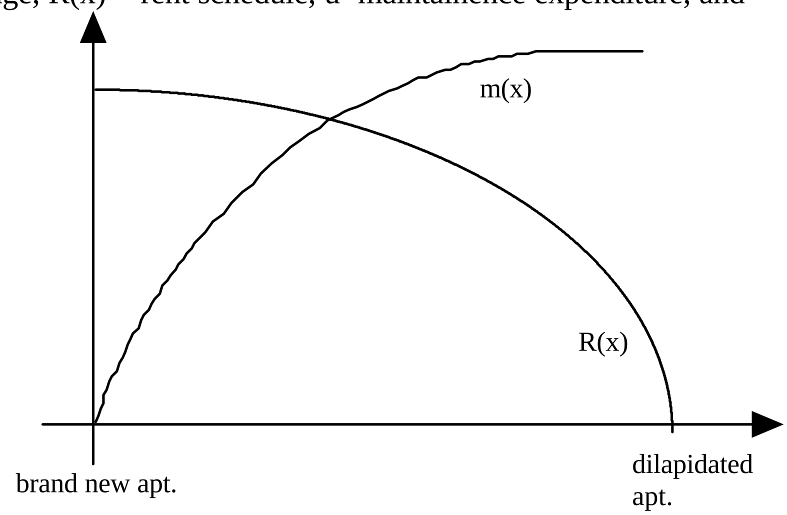
Extension problem: find p(t) such that u(t) is as flat as possible. This is a STACKELBERY LEADER-FOLLOWER game problem.
Other considerations: Block discount, hydro storage

15. Rent Control Problem

$$\dot{x}=1-m(x)u$$

$$J=\int_0^{\cdot} e^{-\delta t}(R(x)-u]dt$$

where x(t)= effective agfe of apartment house, δ=interest rate, m(x) maintainence coeeficient as a function of house age, R(x) = rent schedule, u=maintainence expenditure, and



Extension problem: what is the optimal R(x)?

16. Discrete Event Dynamic Systems vs. Continuous Variable Dynamic Systems (CVDS)

Examples: Flexible Manufacturing Systems, the Internet, The National Air Space (NAS) Traffic Control Systems.
ES205 treats DEDS

Appendix

NOTATIONS AND FACTS ABOUT LINEAR ALGEBRA

1. In general, lower case character denotes vectors and upper case, matrices.

$$x=\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} ; \; A = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ \vdots & & \dots & \vdots \\ a_{m1} & & \dots & a_{mn} \end{bmatrix}$$

except where context makes clear that scalar quantitties are invovled. In ES 202, L and J are usually reserved for scalars. Thus,

$$L(x, u) \equiv L(x_1, \dots, x_n ; u_1, \dots, u_m)$$

is a scalar function of n+m variables.

2. The superscript "T" denotes transpose. Thus,

$$x^T = [\; x_1 \quad \dots \quad x_n \;]$$

3. Simple facts:

i. $(AB)^T=B^TA^T$

ii. The quadratic form $x^TAx \equiv \sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij} x_i x_j = 0$ for any antisymemmtric matrix A, i.e., when $A = -A^T$ .Thus there is no loss of generality to limit the matrix A in any quadratic form to symmetric matrices since any matrix can be decomposed as sum of a symmetric and an antisymmetric matrices. (Note: for any $A=(A^s+A^a)/2$ where $A^s = (A+A^T)/2$ and $A^a= (A-A^T)/2$

iii. A positive definite (semidefinite) matrix A is defined to be any square matrix such that $x^TAx > (\geq) \; 0$ for all $x \neq 0$.. We write A>( )0. Any A> ( ) 0 have only positive (non-negative) eigenvalues

iv. Any square matrix A can be transformed by a similarity tranform into

$$TAT^{-1} = \begin{bmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & & & 0 \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \dots & \lambda_n \end{bmatrix}$$

where the λ_i may be a Jordan canonical block but usually just a scalar eigenvalue. We can view the transformation matrix T as a coordinate tranformation involving rotation, stretching, and inversion of corrdinate axes in n-space. In this sense, only the eigenvalue really matters. In the linear ordinary differential equation dx/dt =Ax. after coordinate transformation we have n uncoupled first order LODE which have exponential (with possibly complex exponents) as basis solutions.

4. The scalar (inner) product of two vectors are defined as

$$x^T y \equiv \sum_{i=1}^{n} x_i y_i$$

For the case of x=y, we write $x^T x \equiv \sum_{i=1}^{n} x_i^2 \equiv ||x||^2$ which is simply the square length of the vector, a positive number for all x 0. More generally, for A>0, $x^T Ax$ is the generalized squared vector length.
Any matrix $A^T A$ for arbitrary A is at least semidefinite since $x^T A^T Ax \equiv y^T y$ 0

5. A set of n vectors $x_1, \ldots, x_n$ are said to be linearly indepndent if there does NOT exist linear combination such that

$$\sum_{i=1}^{n} \alpha_i x_i = 0 \quad \text{for} \quad \alpha_i \text{ not all zero}$$

6. The following statement about the equation Ax=b for square A are equivalent:
   - i. A^{-1} exist
   - ii. A has rank = n
   - iii the equation has unique solution x for any b 0
   - iv. the column (rows) of A are linearly independent.

7. More generally, if A is pxn p n, then there exist infinite number of solutions for x if A has maximal rank=p. The minimal norm solution, i.e., $||x||^2$ is the smallest, is given by $x = A^T (AA^T)^{-1} b$.

8. A totally coordinate-free and mathematically elegant way of dealing with all of above is by way of linear spaces.