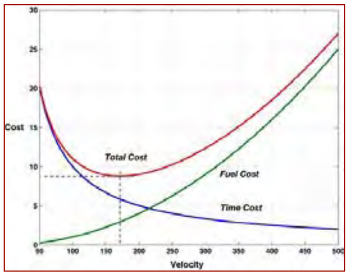
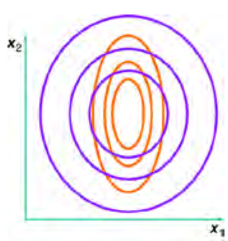


Introduction to Optimization

Robert Stengel
Robotics and Intelligent Systems,
MAE 345, Princeton University, 2015

Optimization problems and criteria
Cost functions
Static optimality conditions
Examples of static optimization



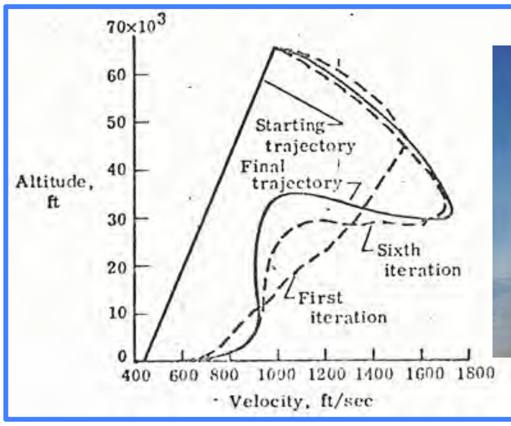
Typical Optimization Problems

- **Minimize** the **probable error** in an estimate of the dynamic state of a system
- **Maximize** the probability of making a **correct decision**
- **Minimize** the **time or energy** required to achieve an objective
- **Minimize** the **regulation error** in a controlled system

- Estimation
- Control

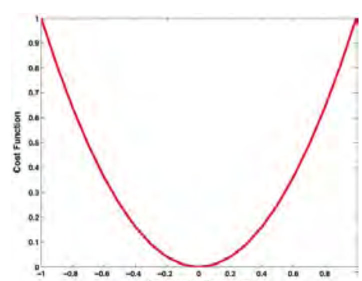
Optimization Implies Choice

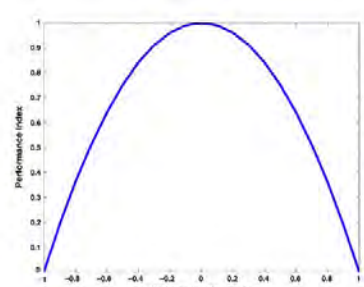
- Choice of best strategy
- Choice of best design parameters
- Choice of best control history
- Choice of best estimate
- Optimization provided by selection of the best control variable



Criteria for Optimization

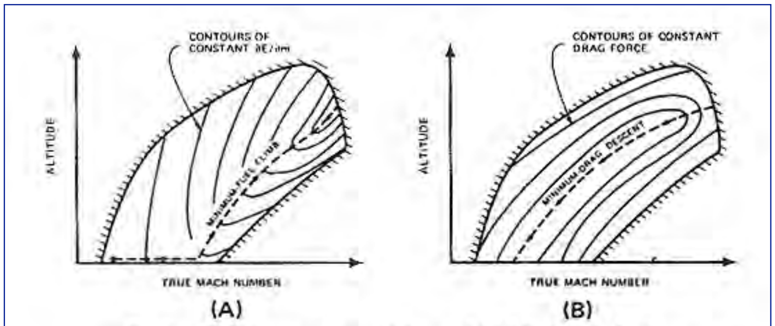
- Names for criteria
  - Figure of merit
  - Performance index
  - Utility function
  - Value function
  - Fitness function
  - Cost function, *J*
    - Optimal cost function = *J**
    - Optimal control = u*
- Different criteria lead to different optimal solutions
- Types of Optimality Criteria
  - Absolute
  - Regulatory
  - Feasible





Minimize Absolute Criteria

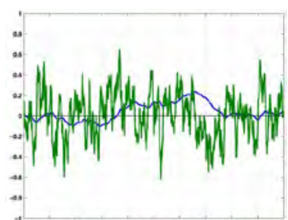
Achieve a specific objective, such as minimizing the required time, fuel, or financial cost to perform a task



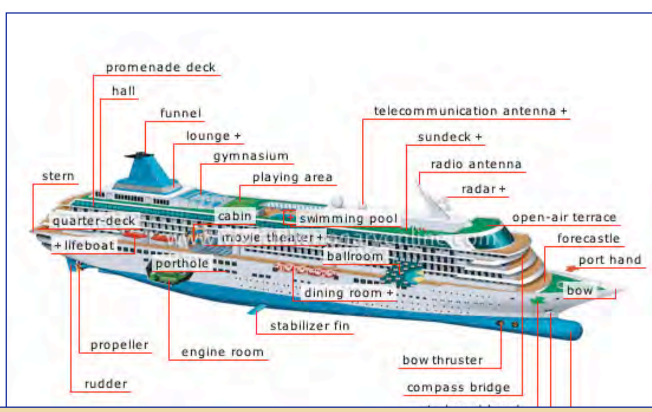
Fuel-optimal energy transitions. (A) Minimum fuel energy climb. (B) Minimum drag energy descent.

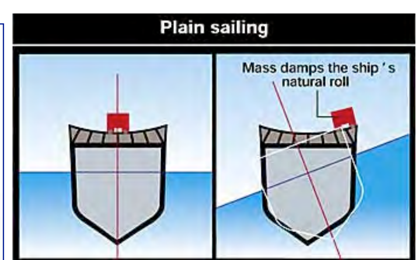
What is the control variable?

Optimal System Regulation



- **Find feedback control gains that minimize tracking error, Δx, in presence of random disturbances**

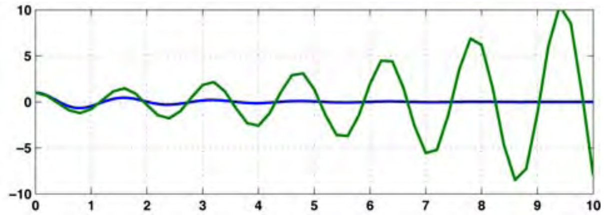


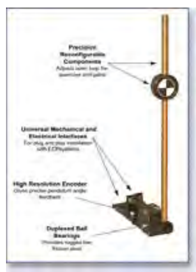


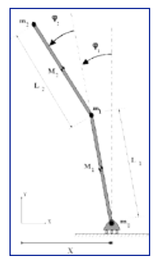
Cruise Ship, Anyone?
http://www.youtube.com/watch?v=bVUFj35BNKM&list=PLCF1F89084A30FBED

Feasible Control Logic

Find feedback control structure that guarantees stability (i.e., that prevents divergence)

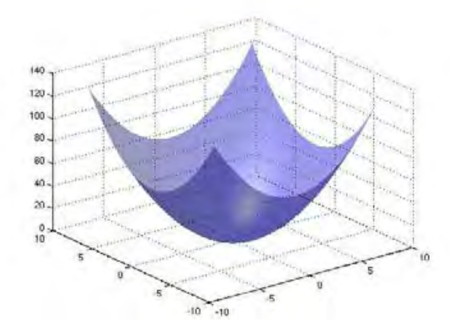




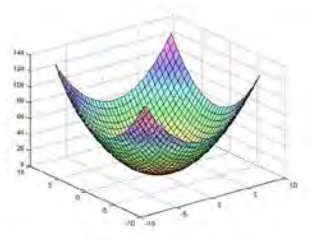


Single Inverted Pendulum	Double Inverted Pendulum
http://www.youtube.com/watch?v=mi-tek7HvZs	*http://www.youtube.com/watch?v=8HDDzKxNMEY*

Desirable Characteristics of a Cost Function

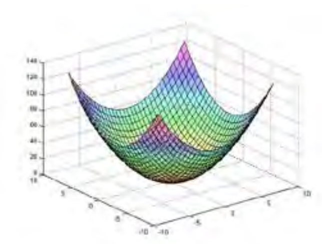


- **Scalar**
- **Clearly defined (preferably unique) maximum or minimum**
  - **Local**
  - **Global**
- **Preferably *positive-definite* (i.e., always a positive number)**

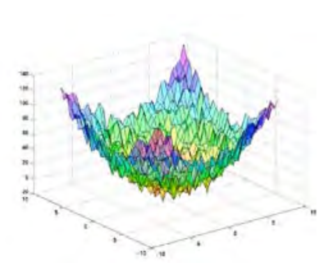


Static vs. Dynamic Optimization

- **Static**
  - Optimal state, **x***, and control, **u***, are fixed, i.e., they do not change over time: $J^* = J(x^*, u^*)$
    - Functional minimization (or maximization)
    - Parameter optimization
- **Dynamic**
  - Optimal state and control vary over time: $J^* = J[x^*(t), u^*(t)]$
    - Optimal trajectory
    - Optimal feedback strategy
- **Optimized cost function, J^*, is a scalar, real number in both cases**



Deterministic vs. Stochastic Optimization

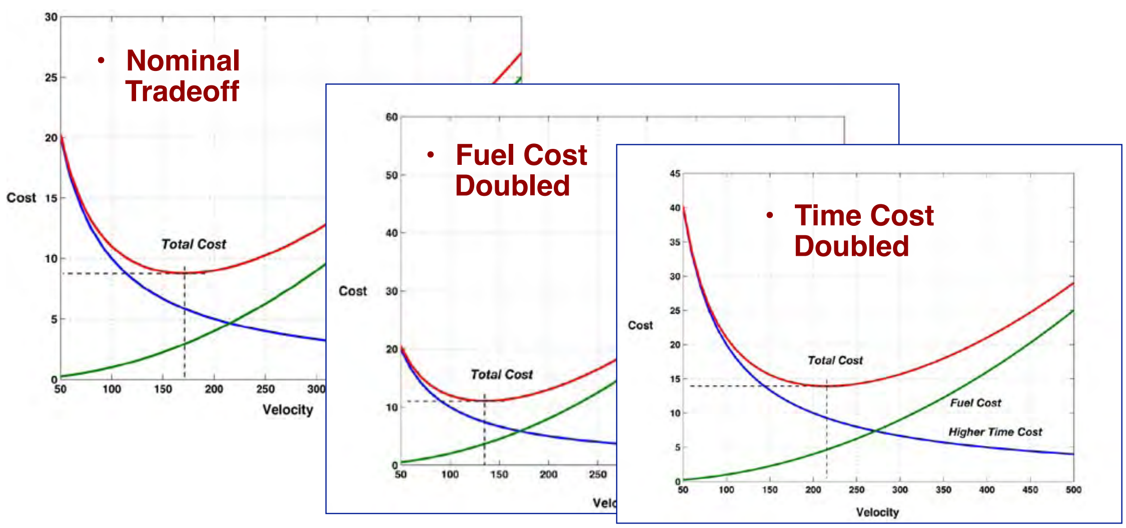


- **Deterministic**
  - System model, parameters, initial conditions, and disturbances are known without error
  - Optimal control operates on the system with certainty
    - $J^* = J(x^*, u^*)$
- **Stochastic**
  - Uncertainty in system model, parameters, initial conditions, disturbances, and resulting cost function
  - Optimal control minimizes the expected value of the cost:
    - *Optimal cost* $= E\{J[x^*, u^*]\}$
- **Cost function is a scalar, real number in both cases**

Cost Function with a Single Control Parameter

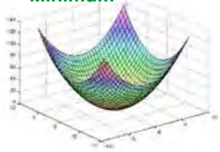
- Tradeoff between two types of cost: Minimum-cost cruising speed
  - Fuel cost proportional to velocity-squared
  - Cost of time inversely proportional to velocity
- Control parameter: Velocity

Tradeoff Between Time- and Fuel-Based Costs

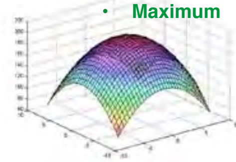


Cost Functions with Two Control Parameters

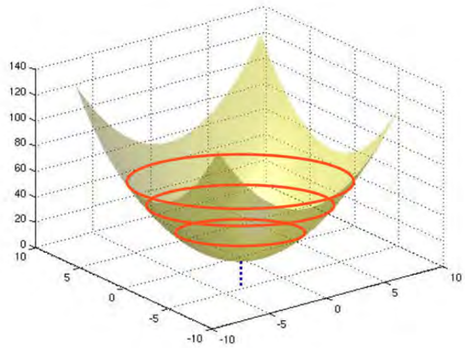
• Minimum



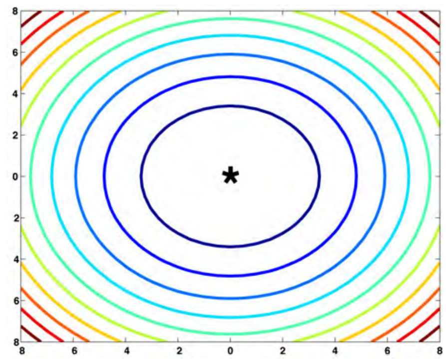
• Maximum



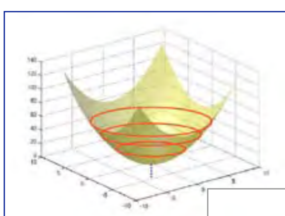
3-D plot of equal-cost contours (iso-contours)



2-D plot of equal-cost contours (iso-contours)

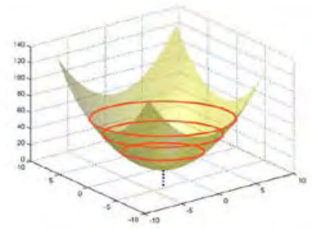


Real-World Topography



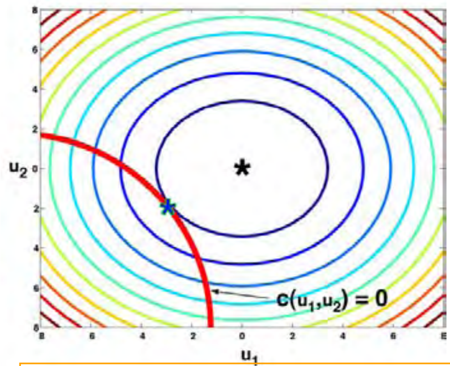
Local vs. global maxima/minima

Robustness of estimates

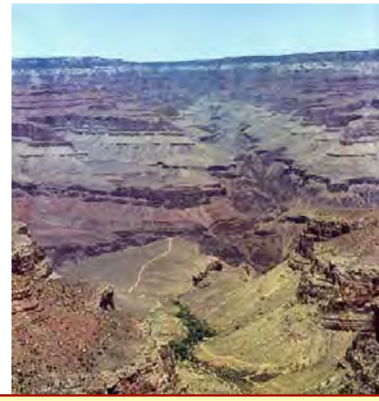


Cost Functions with Equality Constraints

Must stay on the trail



Equality constraint may allow control dimension to be reduced



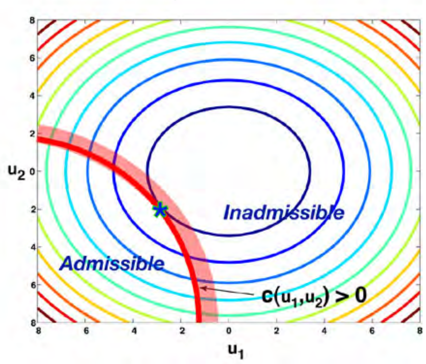
$$c(u_1,u_2)=0 \rightarrow u_2 = fcn(u_1)$$

then

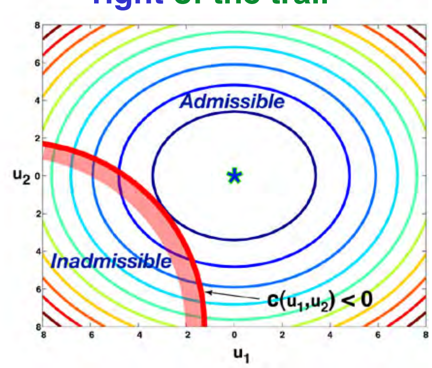
$$J(u_1,u_2)=J\left[u_1, fcn(u_1)\right]=J'(u_1)$$

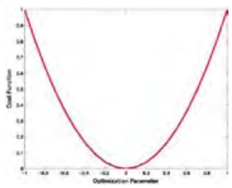
Cost Functions with Inequality Constraints

Must stay to the left of the trail



Must stay to the right of the trail





Necessary Condition for Static Optimality

Single control

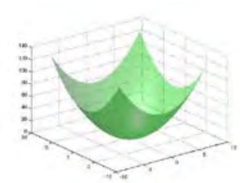
$$\boxed{\left.\frac{dJ}{du}\right|_{u=u^*} = 0}$$

i.e., the slope is zero at the optimum point

Example:

$$J = (u-4)^2$$
$$\frac{dJ}{du} = 2(u-4)$$
$$= 0 \quad when\ u^* = 4$$

Necessary Condition for Static Optimality



Multiple controls

$$\left.\frac{\partial J}{\partial \mathbf{u}}\right|_{\mathbf{u}=\mathbf{u}^*} = \left.\begin{bmatrix} \frac{\partial J}{\partial u_1} & \frac{\partial J}{\partial u_2} & \cdots & \frac{\partial J}{\partial u_m} \end{bmatrix}\right|_{\mathbf{u}=\mathbf{u}^*} = \mathbf{0}$$

Gradient

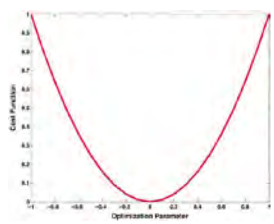
i.e., all slopes are concurrently zero at the optimum point

Example:

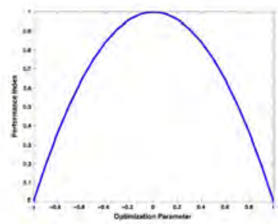
$$J = (u_1-4)^2 + (u_2-8)^2$$
$$dJ/du_1 = 2(u_1-4) = 0 \quad when\ u_1^* = 4$$
$$dJ/du_2 = 2(u_2-8) = 0 \quad when\ u_2^* = 8$$
$$\left.\frac{\partial J}{\partial \mathbf{u}}\right|_{\mathbf{u}=\mathbf{u}^*} = \left.\begin{bmatrix} \frac{\partial J}{\partial u_1} & \frac{\partial J}{\partial u_2} \end{bmatrix}\right|_{\mathbf{u}=\mathbf{u}^*=\begin{bmatrix} 4 \\ 8 \end{bmatrix}} = \begin{bmatrix} 0 & 0 \end{bmatrix}$$

… But the Slope can be Zero for More than One Reason

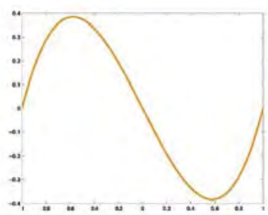
Minimum



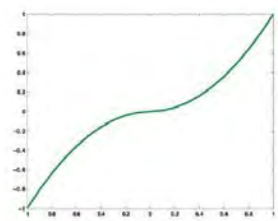
Maximum

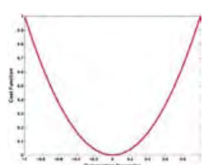


Either



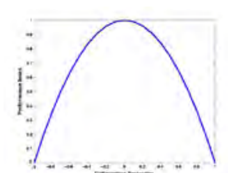
Inflection Point





Sufficient Condition for Static Optimum

- Single control



Minimum

Satisfy necessary condition plus

$$\boxed{\left.\frac{d^2 J}{du^2}\right|_{u=u^*} > 0}$$

i.e., curvature is positive at optimum

Example:

$$J = (u-4)^2$$

$$\frac{dJ}{du} = 2(u-4)$$

$$\frac{d^2 J}{du^2} = 2 > 0$$

Maximum

Satisfy necessary condition plus

$$\boxed{\left.\frac{d^2 J}{du^2}\right|_{u=u^*} < 0}$$

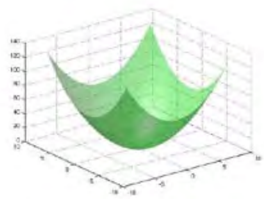
i.e., curvature is negative at optimum

Example:

$$J = -(u-4)^2$$

$$\frac{dJ}{du} = -2(u-4)$$

$$\frac{d^2 J}{du^2} = -2 < 0$$



Sufficient Condition for Static Minimum

Multiple controls

- **Satisfy necessary condition**
  - plus

$$\left.\frac{\partial J}{\partial \mathbf{u}}\right|_{\mathbf{u}=\mathbf{u}^*} = \left[\frac{\partial J}{\partial u_1} \quad \frac{\partial J}{\partial u_2} \quad \cdots \quad \frac{\partial J}{\partial u_m}\right]_{\mathbf{u}=\mathbf{u}^*} = \mathbf{0}$$

$$\left.\frac{\partial^2 J}{\partial \mathbf{u}^2}\right|_{\mathbf{u}=\mathbf{u}^*} = \left[\begin{array}{cccc} \frac{\partial^2 J}{\partial u_1^2} & \frac{\partial^2 J}{\partial u_1 \partial u_2} & \cdots & \frac{\partial^2 J}{\partial u_1 \partial u_m} \\ \frac{\partial^2 J}{\partial u_2 \partial u_1} & \frac{\partial^2 J}{\partial u_2^2} & \cdots & \frac{\partial^2 J}{\partial u_2 \partial u_m} \\ \cdots & \cdots & \cdots & \cdots \\ \frac{\partial^2 J}{\partial u_m \partial u_1} & \frac{\partial^2 J}{\partial u_2 \partial u_m} & \cdots & \frac{\partial^2 J}{\partial u_m^2} \end{array}\right]_{\mathbf{u}=\mathbf{u}^*} > \mathbf{0}$$

Hessian matrix

- … what does it mean for a matrix to be "greater than zero"?

$$\frac{\partial^2 J}{\partial \mathbf{u}^2} \triangleq \mathbf{Q} > \mathbf{0}$$ **if Its Quadratic Form, $\mathbf{x}^T\mathbf{Q}\mathbf{x}$, is Greater than Zero**

$$\mathbf{x}^T\mathbf{Q}\mathbf{x} \triangleq \text{ Quadratic form}$$

$\mathbf{Q}$: **Defining matrix** of the quadratic form

$$\left[(1\times n)(n\times n)(n\times 1)\right] = \left[(1\times 1)\right]$$

- dim(Q) = n x n
- Q is symmetric
- $x^T Qx$ is a scalar

Quadratic Form of Q is Positive* if Q is Positive Definite

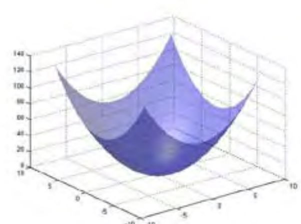
- **Q is positive-definite if**
  - All leading principal minor determinants are positive
  - All eigenvalues are real and positive
    - 3 x 3 example

$$\mathbf{Q} = \begin{bmatrix} q_{11} & q_{12} & q_{13} \\ q_{21} & q_{22} & q_{23} \\ q_{31} & q_{32} & q_{33} \end{bmatrix}$$

$$q_{11} > 0, \quad \begin{vmatrix} q_{11} & q_{12} \\ q_{21} & q_{22} \end{vmatrix} > 0, \quad \begin{vmatrix} q_{11} & q_{12} & q_{13} \\ q_{21} & q_{22} & q_{23} \\ q_{31} & q_{32} & q_{33} \end{vmatrix} > 0$$

* *except at* **x = 0**

$$\det(s\mathbf{I} - \mathbf{Q}) = (s - \lambda_1)(s - \lambda_2)(s - \lambda_3)$$
$$\lambda_1, \lambda_2, \lambda_3 \geq 0$$



Minimized Cost Function, *J**

- **Gradient** is zero at the minimum
- **Hessian matrix** is positive-definite at the minimum
- Expand the cost in a *Taylor series*

$$J(\mathbf{u}^* + \Delta\mathbf{u}) \approx J(\mathbf{u}^*) + \Delta J(\mathbf{u}^*) + \Delta^2 J(\mathbf{u}^*) + \ldots$$

$$\Delta J(\mathbf{u}^*) = \Delta\mathbf{u}^T \left.\frac{\partial J}{\partial \mathbf{u}}\right|_{\mathbf{u}=\mathbf{u}^*} = 0$$

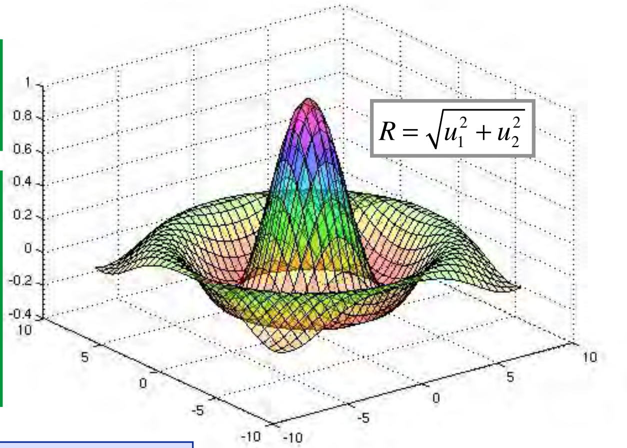
$$\Delta^2 J(\mathbf{u}^*) = \frac{1}{2}\Delta\mathbf{u}^T \left[\left.\frac{\partial^2 J}{\partial \mathbf{u}^2}\right|_{\mathbf{u}=\mathbf{u}^*}\right]\Delta\mathbf{u} \geq 0$$

- **First variation** is zero at the minimum
- **Second variation** is positive at the minimum

How Many Maxima/Minima does the "Mexican Hat" [z = sinc R = (sin R)/R] Have?

$$\left.\frac{\partial J}{\partial \mathbf{u}}\right|_{\mathbf{u}=\mathbf{u}^*} = \left.\begin{bmatrix} \dfrac{\partial J}{\partial u_1} & \dfrac{\partial J}{\partial u_2} & \dots & \dfrac{\partial J}{\partial u_m} \end{bmatrix}\right|_{\mathbf{u}=\mathbf{u}^*} = \mathbf{0}$$

$$\left.\frac{\partial^2 J}{\partial \mathbf{u}^2}\right|_{\mathbf{u}=\mathbf{u}^*} = \left.\begin{bmatrix} \dfrac{\partial^2 J}{\partial u_1^2} & \dfrac{\partial^2 J}{\partial u_1 \partial u_2} & \dots & \dfrac{\partial^2 J}{\partial u_1 \partial u_m} \\ \dfrac{\partial^2 J}{\partial u_2 \partial u_1} & \dfrac{\partial^2 J}{\partial u_2^2} & \dots & \dfrac{\partial^2 J}{\partial u_2 \partial u_m} \\ \dots & \dots & \dots & \dots \\ \dfrac{\partial^2 J}{\partial u_m \partial u_1} & \dfrac{\partial^2 J}{\partial u_2 \partial u_m} & \dots & \dfrac{\partial^2 J}{\partial u_m^2} \end{bmatrix}\right|_{\mathbf{u}=\mathbf{u}^*} \gtrless \mathbf{0}$$

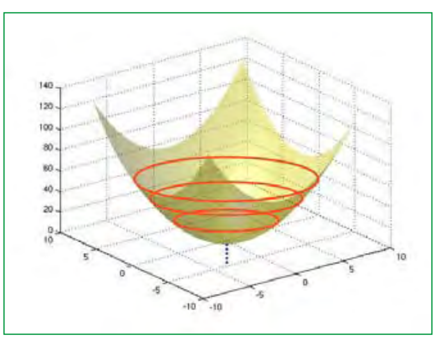


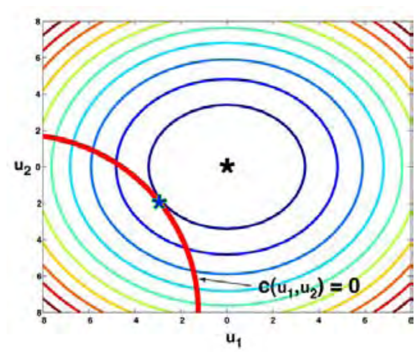
One maximum

Wolfram Alpha
maximize(sinc(sqrt(x^2+y^2)))

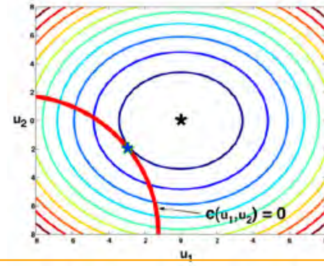
Static Cost Functions with Equality Constraints

- **Minimize $J(u')$, subject to c(u') = 0**
  - **dim(c) = [n x 1]**
  - **dim(u') = [(m + n) x 1]**





Two Approaches to Static Optimization with a Constraint



$$\mathbf{u}' = \begin{bmatrix} u_1 \\ u_2 \end{bmatrix}$$

1. Use constraint to reduce control dimension
2. Augment the cost function to recognize the constraint

$$\textit{Example}: \quad \min_{u_1,u_2} J \quad \textit{subject to}$$

$$c(\mathbf{u}') = c(u_1,u_2) = 0 \rightarrow u_2 = fcn(u_1)$$

then

$$J(\mathbf{u}') = J(u_1,u_2) = J[u_1, fcn(u_1)] = J'(u_1)$$

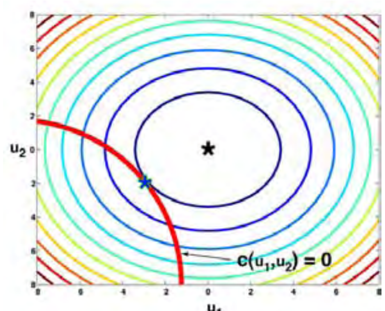
$$J_A(\mathbf{u}') = J(\mathbf{u}') + \boldsymbol{\lambda}^T \mathbf{c}(\mathbf{u}')$$

$\boldsymbol{\lambda}$, an unknown constant

$\boldsymbol{\lambda}$ has the same dimension as the constraint

$$\dim(\boldsymbol{\lambda}) = \dim(\mathbf{c}) = n \times 1$$

Solution:
First Approach



Cost function

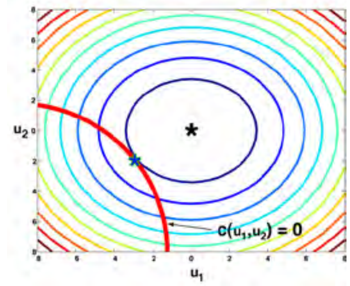
$$J = u_1^{\;2} - 2u_1u_2 + 3u_2^{\;2} - 40$$

Constraint

$$c = u_2 - u_1 - 2 = 0$$

$$\therefore u_2 = u_1 + 2$$

Solution Example: Reduced Control Dimension



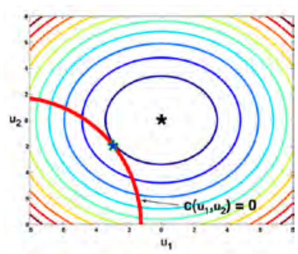
Cost function and gradient with substitution

$$
\begin{aligned}
J &= u_1^2 - 2u_1u_2 + 3u_2^2 - 40 \\
&= u_1^2 - 2u_1(u_1+2) + 3(u_1+2)^2 - 40 \\
&= 2u_1^2 + 8u_1 - 28
\end{aligned}
$$

$$\frac{\partial J}{\partial u_1} = 4u_1 + 8 = 0$$

Optimal solution

$$
\begin{aligned}
u_1^* &= -2 \\
u_2^* &= 0 \\
J^* &= -36
\end{aligned}
$$



Solution: Second Approach

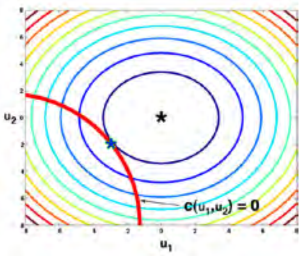
- Partition **u'** into a state, **x**, and a control, **u**, such that
  - dim(x) = [n x 1]
  - dim(u) = [m x 1]

$$\mathbf{u}' = \begin{bmatrix} \mathbf{x} \\ \mathbf{u} \end{bmatrix}$$

- Add constraint to the cost function, weighted by Lagrange multiplier, **λ**
- **c** is required to be zero when J_A is a minimum

$$
\begin{aligned}
J_A(\mathbf{u}') &= J(\mathbf{u}') + \boldsymbol{\lambda}^T \mathbf{c}(\mathbf{u}') \\
J_A(\mathbf{x},\mathbf{u}) &= J(\mathbf{x},\mathbf{u}) + \boldsymbol{\lambda}^T \mathbf{c}(\mathbf{x},\mathbf{u})
\end{aligned}
$$

$$\mathbf{c}(\mathbf{u}') = \mathbf{c}\begin{pmatrix} \mathbf{x} \\ \mathbf{u} \end{pmatrix} = \mathbf{0}$$



Solution: Adjoin Constraint with Lagrange Multiplier

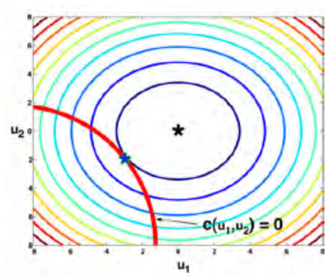
Gradient with respect to x, u, and λ is zero at the optimum point

$$\frac{\partial J_A}{\partial \mathbf{x}} = \frac{\partial J}{\partial \mathbf{x}} + \boldsymbol{\lambda}^T \frac{\partial \mathbf{c}}{\partial \mathbf{x}} = \mathbf{0}$$

$$\frac{\partial J_A}{\partial \mathbf{u}} = \frac{\partial J}{\partial \mathbf{u}} + \boldsymbol{\lambda}^T \frac{\partial \mathbf{c}}{\partial \mathbf{u}} = \mathbf{0}$$

$$\frac{\partial J_A}{\partial \boldsymbol{\lambda}} = \mathbf{c} = \mathbf{0}$$

Simultaneous Solutions for State and Control



- ($2n + m$) values must be found (x, λ, u)
- Use first equation to find form of optimizing Lagrange multiplier (n scalar equations)
- Second and third equations provide ($n + m$) scalar equations that specify the state and control

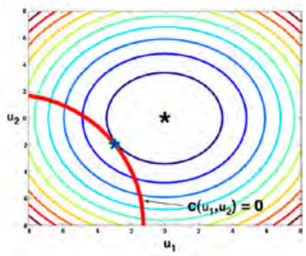
$$\boldsymbol{\lambda}^{*T} = -\frac{\partial J}{\partial \mathbf{x}}\left(\frac{\partial \mathbf{c}}{\partial \mathbf{x}}\right)^{-1}$$

$$\boldsymbol{\lambda}^{*} = -\left[\left(\frac{\partial \mathbf{c}}{\partial \mathbf{x}}\right)^{-1}\right]^T \left(\frac{\partial J}{\partial \mathbf{x}}\right)^T$$

$$\frac{\partial J}{\partial \mathbf{u}} + \boldsymbol{\lambda}^{*T} \frac{\partial \mathbf{c}}{\partial \mathbf{u}} = \mathbf{0}$$

$$\frac{\partial J}{\partial \mathbf{u}} - \frac{\partial J}{\partial \mathbf{x}}\left(\frac{\partial \mathbf{c}}{\partial \mathbf{x}}\right)^{-1} \frac{\partial \mathbf{c}}{\partial \mathbf{u}} = \mathbf{0}$$

$$\mathbf{c}(\mathbf{x}, \mathbf{u}) = \mathbf{0}$$



Solution Example: Second Approach

Cost function

$$\boxed{J = u^2 - 2xu + 3x^2 - 40}$$

Constraint

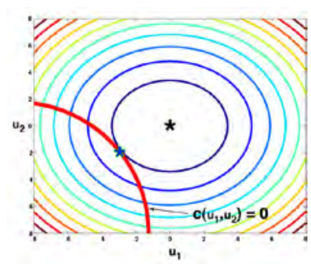
$$\boxed{c = x - u - 2 = 0}$$

Partial derivatives

$$\frac{\partial J}{\partial x} = -2u + 6x \qquad \frac{\partial c}{\partial x} = 1$$

$$\frac{\partial J}{\partial u} = 2u - 2x \qquad \frac{\partial c}{\partial u} = -1$$

Solution Example: Second Approach



- **From first equation** $\quad \boxed{\lambda^* = 2u - 6x}$
- **From second equation** $$(2u - 2x) + (2u - 6x)(-1)$$ $$\therefore x = 0$$
- **From constraint** $\quad \boxed{u = -2}$
  - **Optimal solution**

$$x^* = 0$$
$$u^* = -2$$
$$J^* = -36$$

Next Time:
Numerical Optimization