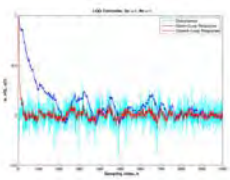
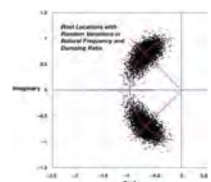
Stochastic Control

Robert Stengel
Robotics and Intelligent Systems, MAE 345,
Princeton University, 2017

Learning Objectives

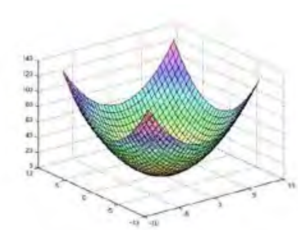
- **Markov Process**
- **Overview of the Linear-Quadratic-Gaussian (LQG) Regulator**
- **Introduction to Stochastic Robust Control Laws**



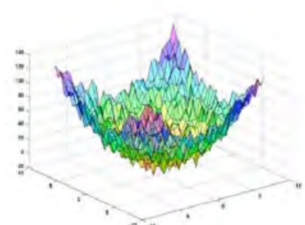


Deterministic vs. Stochastic Optimal Control

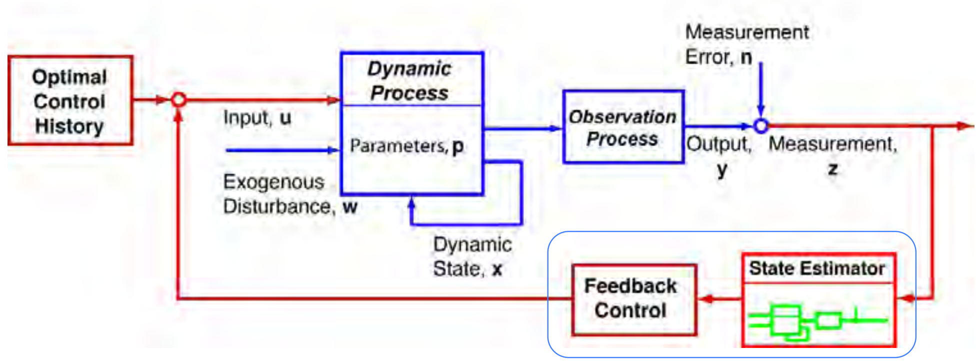
- **Deterministic control**
  - **Known dynamic process**
    - **precise input**
    - **precise initial condition**
    - **precise measurement**
  - **Optimal control minimizes $J^* = J(x^*, u^*)$**



- **Stochastic control**
  - **Known dynamic process**
    - **unknown input**
    - **imprecise initial condition**
    - **imprecise or incomplete measurement**
  - **Optimal control minimizes $E\{J[x^*, u^*]\}$**



Linear-Quadratic-Gaussian (LQG) Control of a Dynamic Process



$$\mathbf{z}(t) = \mathbf{H}\mathbf{x}(t) + \mathbf{n}(t)$$

$$\dot{\hat{\mathbf{x}}}(t) = \mathbf{F}(t)\hat{\mathbf{x}}(t) + \mathbf{G}(t)\mathbf{u}(t) + \mathbf{K}(t)\left[\mathbf{z}(t) - \mathbf{H}\hat{\mathbf{x}}(t)\right]$$

$$\mathbf{u}(t) = -\mathbf{C}(t)\,\hat{\mathbf{x}}(t) + \mathbf{u}_{command}$$

3

Linear-Quadratic (LQ) Control Equations

(Continuous-Time Model)

Open-Loop System State Dynamics

$$\dot{\mathbf{x}}(t) = \mathbf{F}\mathbf{x}(t) + \mathbf{G}\mathbf{u}(t)$$

Control Law

$$\mathbf{u}(t) = -\mathbf{C}\mathbf{x}(t) + \mathbf{u}_{command}$$

Closed-Loop System State Dynamics

$$\dot{\mathbf{x}}(t) = (\mathbf{F} - \mathbf{G}\mathbf{C})\mathbf{x}(t) + \mathbf{G}\mathbf{C}\mathbf{u}_{command}$$

Characteristic Equation

$$\left|s\mathbf{I} - (\mathbf{F} - \mathbf{G}\mathbf{C})\right| = 0$$

How many eigenvalues?

n

Stable or unstable?

Stable, with correct design criteria, F, and G

4

Linear-Quadratic-Gaussian (LQG) Control Equations
(Continuous-Time Model)

Open-Loop System State Dynamics and Measurement

$$\dot{\mathbf{x}}(t) = \mathbf{F}\mathbf{x}(t) + \mathbf{G}\mathbf{u}(t) + \mathbf{L}\mathbf{w}(t)$$
$$\mathbf{z}(t) = \mathbf{H}\mathbf{x}(t) + \mathbf{n}(t)$$

State Estimate

$$\dot{\hat{\mathbf{x}}}(t) = \mathbf{F}\hat{\mathbf{x}}(t) + \mathbf{G}\mathbf{u}(t) + \mathbf{K}\left[\mathbf{z}(t) - \mathbf{H}\hat{\mathbf{x}}(t)\right]$$

Control Law

$$\mathbf{u}(t) = -\mathbf{C}\hat{\mathbf{x}}(t) + \mathbf{u}_{command}$$

Linear-Quadratic-Gaussian (LQG) Control Equations
(Continuous-Time Model)

Closed-Loop System State and Estimate Dynamics
(neglect command)

$$\dot{\mathbf{x}}(t) = \mathbf{F}\mathbf{x}(t) + \mathbf{G}\left[-\mathbf{C}\hat{\mathbf{x}}(t)\right] + \mathbf{L}\mathbf{w}(t)$$

$$\dot{\hat{\mathbf{x}}}(t) = \mathbf{F}\hat{\mathbf{x}}(t) + \mathbf{G}\left[-\mathbf{C}\hat{\mathbf{x}}(t)\right] + \mathbf{K}\left[\mathbf{z}(t) - \mathbf{H}\hat{\mathbf{x}}(t)\right]$$

How many eigenvalues? ***2n***

Stable or unstable? ***TBD***

LQG Separation Property

Optimal estimation algorithm does not depend on the optimal control algorithm

$$\mathbf{K}(t) = \mathbf{P}(t)\mathbf{H}^T\mathbf{N}^{-1}(t)$$

$$\dot{\mathbf{P}}(t) = \mathbf{F}(t)\mathbf{P}(t) + \mathbf{P}(t)\mathbf{F}^T(t) + \mathbf{L}(t)\mathbf{W}(t)\mathbf{L}^T(t) - \mathbf{P}(t)\mathbf{H}^T\mathbf{N}^{-1}(t)\mathbf{H}\mathbf{P}(t)$$

Optimal control algorithm does not depend on the optimal estimation algorithm

$$\mathbf{C}(t) = \mathbf{R}^{-1}(t)\mathbf{G}^T(t)\mathbf{S}(t)$$

$$\dot{\mathbf{S}}(t) = -\mathbf{Q}(t) - \mathbf{F}(t)^T\mathbf{S}(t) - \mathbf{S}(t)\mathbf{F}(t) + \mathbf{S}(t)\mathbf{G}(t)\mathbf{R}^{-1}(t)\mathbf{G}^T(t)\mathbf{S}(t)$$

LQG Certainty Equivalence

Stochastic feedback control is computed from optimal estimate of the state

$$\mathbf{u}^*(t) = -\mathbf{R}^{-1}\mathbf{G}^T(t)\mathbf{S}(t)\hat{\mathbf{x}}(t) = -\mathbf{C}(t)\,\hat{\mathbf{x}}(t)$$

Stochastic feedback control law is the same as the deterministic control law

$$\mathbf{u}^*(t) = -\mathbf{R}^{-1}\mathbf{G}^T(t)\mathbf{S}(t)\mathbf{x}(t) = -\mathbf{C}(t)\,\mathbf{x}(t)$$

Asymptotic Stability of the LQG Regulator

(with no parameter uncertainty)

System Equations with Continuous-Time LQG Control

With perfect knowledge of the system

$$\dot{\mathbf{x}}(t) = \mathbf{F}\mathbf{x}(t) + \mathbf{G}\left[-\mathbf{C}\hat{\mathbf{x}}(t)\right] + \mathbf{L}\mathbf{w}(t)$$

$$\dot{\hat{\mathbf{x}}}(t) = \mathbf{F}\hat{\mathbf{x}}(t) + \mathbf{G}\left[-\mathbf{C}\hat{\mathbf{x}}(t)\right] + \mathbf{K}\left[\mathbf{z}(t) - \mathbf{H}\hat{\mathbf{x}}(t)\right]$$

State estimate error

$$\boldsymbol{\varepsilon}(t) \triangleq \mathbf{x}(t) - \hat{\mathbf{x}}(t)$$

State estimate error dynamics

$$\dot{\boldsymbol{\varepsilon}}(t) = (\mathbf{F} - \mathbf{K}\mathbf{H})\boldsymbol{\varepsilon}(t) + \mathbf{L}\mathbf{w}(t) - \mathbf{K}\mathbf{n}(t)$$

Control-Loop and Estimator Eigenvalues are Uncoupled

$$\begin{bmatrix} \dot{\mathbf{x}}(t) \\ \dot{\boldsymbol{\varepsilon}}(t) \end{bmatrix} = \begin{bmatrix} (\mathbf{F}-\mathbf{GC}) & \mathbf{GC} \\ \mathbf{0} & (\mathbf{F}-\mathbf{KH}) \end{bmatrix} \begin{bmatrix} \mathbf{x}(t) \\ \boldsymbol{\varepsilon}(t) \end{bmatrix} + \begin{bmatrix} \mathbf{L} & \mathbf{0} \\ \mathbf{L} & -\mathbf{K} \end{bmatrix} \begin{bmatrix} \mathbf{w}(t) \\ \mathbf{n}(t) \end{bmatrix}$$

Upper-block-triangular stability matrix

LQG system is stable because

(F – GC) is stable

(F – KH) is stable

Estimate error affects state response

$$\dot{\mathbf{x}}(t) = (\mathbf{F}-\mathbf{GC})\mathbf{x}(t) + \mathbf{GC}\boldsymbol{\varepsilon}(t) + \mathbf{Lw}(t)$$

Actual state does not affect error response

Disturbance affects both equally

Discrete-Time LQG Controller

Kalman filter produces state estimate

$$\hat{\mathbf{x}}_k(-) = \boldsymbol{\Phi}\hat{\mathbf{x}}_{k-1}(+) - \boldsymbol{\Gamma}\mathbf{C}_{k-1}\hat{\mathbf{x}}_{k-1}(+)$$

$$\hat{\mathbf{x}}_k(+) = \hat{\mathbf{x}}_k(-) + \mathbf{K}_k\left[\mathbf{z}_k - \mathbf{H}\hat{\mathbf{x}}_k(-)\right]$$

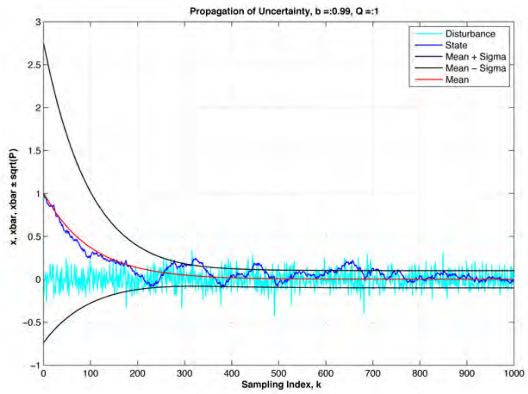
Closed-loop system uses state estimate for feedback control ($u_{command}$ = 0)

$$\mathbf{u}_k = -\mathbf{C}_k\hat{\mathbf{x}}_k(+)$$

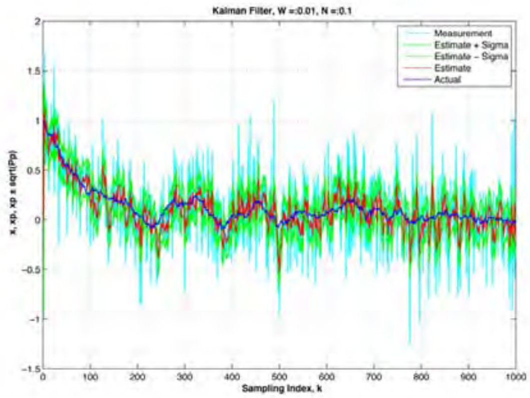
$$\mathbf{x}_{k+1}(-) = \boldsymbol{\Phi}\mathbf{x}_k(-) - \boldsymbol{\Gamma}\mathbf{C}_k\hat{\mathbf{x}}_k(+)$$

Response of Discrete-Time 1st-Order System to Disturbance, and Kalman Filter Estimate from Noisy Measurement

Propagation of Uncertainty

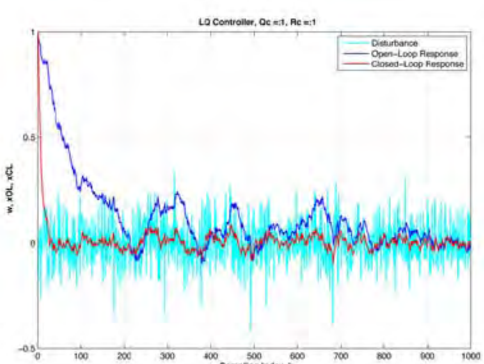


Kalman Filter, Uncontrolled System

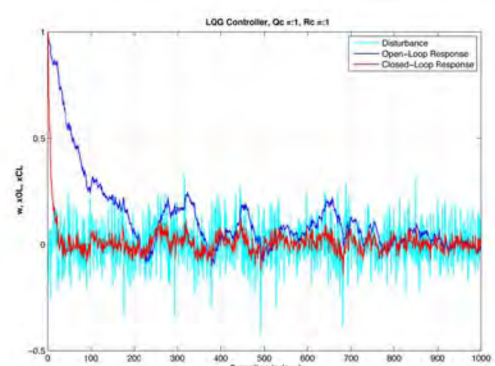


Comparison of 1st-Order Discrete-Time LQ and LQG Control Response

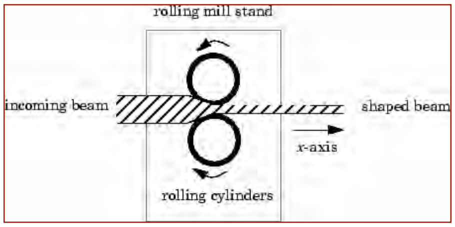
Linear-Quadratic Control with Noise-free Measurement



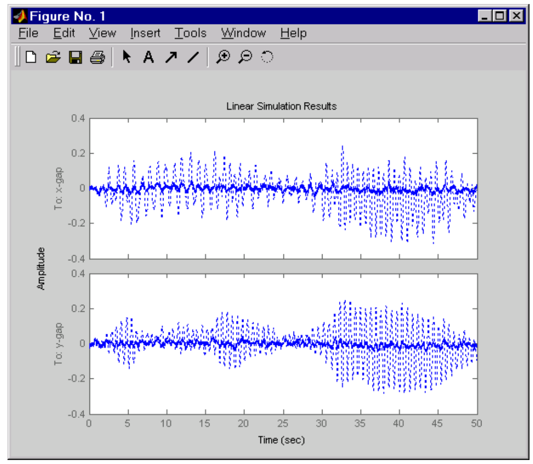
Linear-Quadratic-Gaussian Control with Noisy Measurement



MATLAB Demo: LQG Rolling Mill Control System Design Example



Open- and Closed-Loop Response



- **Maintain desired thickness of shaped beam**
- **Account for random**
  - **variations in thickness/ hardness of incoming beam**
  - **eccentricity in rolling cylinders**
  - **measurement errors**

http://www.mathworks.com/help/control/ug/lqg-regulation-rolling-mill-example.html

15

Robust Stochastic Control

Stochastic, Robust, and Adaptive Control

- Stochastic controller
  - minimize response to random initial conditions, disturbances, and measurement errors
  - perfect knowledge of the plant
- Robust controller
  - fixed gains and structure
  - minimize likelihood of instability or unsatisfactory performance due to parameter uncertainty in the plant
- Adaptive controller
  - variable gains and/or structure
  - minimize likelihood of instability or unsatisfactory performance due to plant parameter uncertainty, disturbances, and measurement errors

Practical controller may have elements of all three

Robust Control System Design

- Make closed-loop response insensitive to **plant parameter variations**
- Robust controller
  - Fixed gains and structure
  - Minimize likelihood of instability
  - Minimize likelihood of unsatisfactory performance

Probabilistic Robust Control Design

- **Design a fixed-parameter controller for stochastic robustness**
- **Monte Carlo Evaluation of competing designs**
- **Genetic Algorithm or Simulated Annealing search for best design**

Representations of Uncertainty

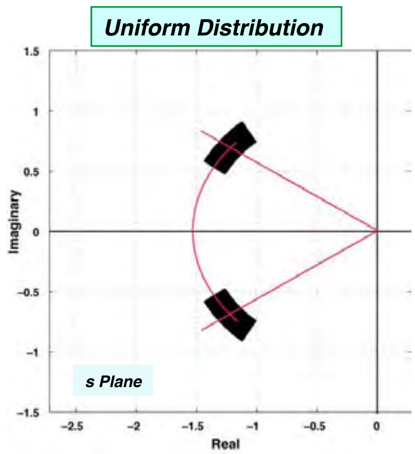
Characteristic equation of the uncontrolled system

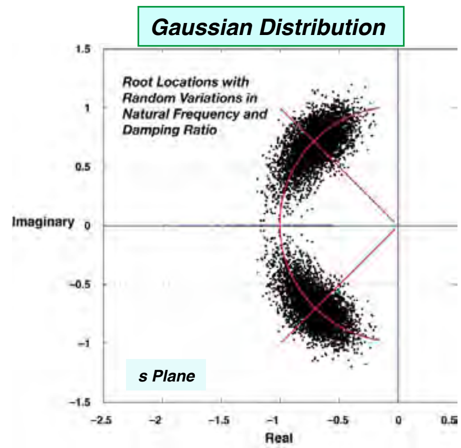
$$
\begin{aligned}
|s\mathbf{I}-\mathbf{F}| &= \det(s\mathbf{I}-\mathbf{F}) \triangleq \\
\Delta(s) &= s^n + a_{n-1}s^{n-1} + \ldots + a_1 s + a_0 \\
&= (s-\lambda_1)(s-\lambda_2)(\ldots)(s-\lambda_n) = 0
\end{aligned}
$$

- **Uncertainty can be expressed in**
  - **Elements of F**
  - **Coefficients of $\Delta(s)$**
  - **Eigenvalues of F**

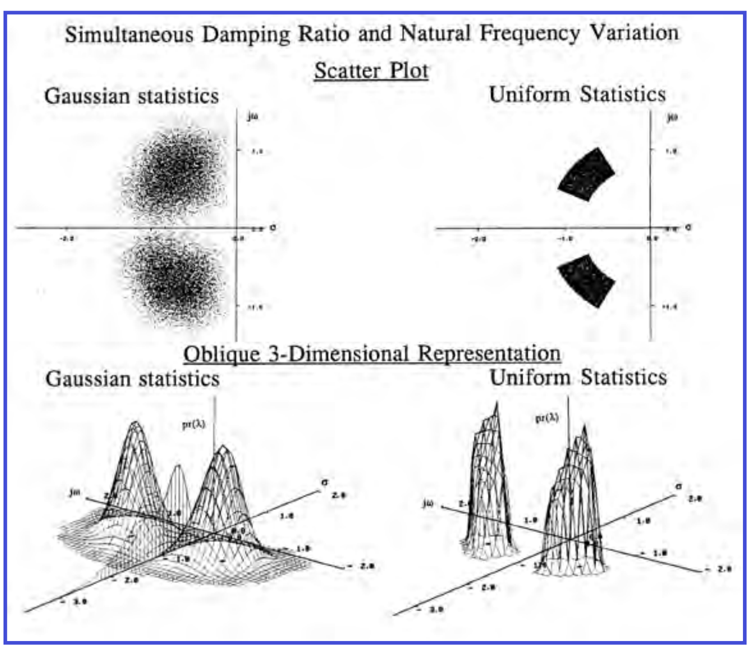
Root Locations for an Uncertain 2nd-Order System

- **Variation may be represented by**
  - **Worst-case, e.g., Upper/lower bounds of uniform distribution**
  - **Probability, e.g., Gaussian distribution**



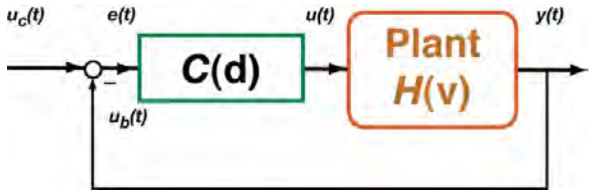


"3-D" Stochastic Root Loci for 2nd-Order Example



- **Root distributions are nonlinear functions of parameter distributions**
- **Unbounded parameter distributions always lead to non-zero probability of instability**
- **Bounded distributions may be guaranteed to be stable**

Probability of Satisfying a Design Metric



$$Pr(\mathbf{d},\mathbf{v}) \approx \frac{1}{N}\sum_{i=1}^{N} e\left[C(\mathbf{d}), H(\mathbf{v})\right]$$

- **Probability of satisfying a design metric**
  - **d:** **Control design parameter vector [e.g., SA, GA, …]**
  - **v:** **Uncertain plant parameter vector [e.g., RNG]**
  - ***e*:** **Binary indicator, e.g.,**
    0: satisfactory 1: unsatisfactory
  - ***H*(v):** **Plant**
  - ***C*(d):** **Controller (Compensator)**

Design Control System to Minimize Probability of Instability

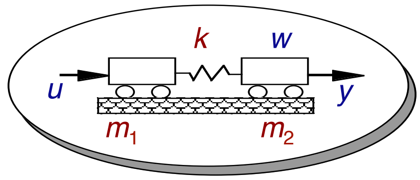
- **Characteristic equation of the closed-loop system**

$$\Delta_{closed-loop}(s) = \left|s\mathbf{I} - \left[\mathbf{F}(\mathbf{v}) - \mathbf{G}(\mathbf{v})\mathbf{C}(\mathbf{d})\right]\right|$$
$$= \left[(s-\lambda_1)(s-\lambda_2)(\ldots)(s-\lambda_n)\right]_{closed-loop} = 0$$

- **Monte Carlo evaluation of probability of instability with uncertain plant parameters**
- **Minimize probability of instability using numerical search of control parameters**

$$\min_{\mathbf{d}}\left\{\Pr\left[\mathrm{Re}\left(\lambda_i, i=1,n\right)\right] > 0\right\}$$

Control Design Example*



- **Challenge: Design a feedback compensator for a 4th-order spring-mass system ("the plant") whose parameters are bounded but unknown**
  - **Minimize the likelihood of instability**
  - **Satisfy a settling time requirement**
  - **Don't use too much control**

** 1990 American Control Conference Robust Control Benchmark Problem*

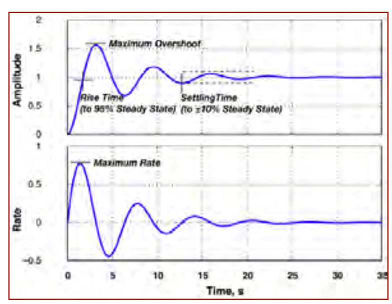
Design Cost Function

- **Probability of Instability, Pr_i**
  - **e_i = 1 (unstable) or 0 (stable)**
- **Probability of Settling Time Exceedance, Pr_{ts}**
  - **e_{ts} = 1 (exceeded) or 0 (not exceeded)**
- **Probability of Control Limit Exceedance, Pr_u**
  - **e_u = 1 (exceeded) or 0 (not exceeded)**
- **Each metric has a binomial distribution**

$$\text{pr}(x) = \frac{n!}{k!(n-k)!} p(x)^k \left[1 - p(x)\right]^{n-k} \triangleq \binom{n}{k} p(x)^k \left[1 - p(x)\right]^{n-k}$$

$=$ probability of exactly k successes in n trials, in $(0,1)$

$\sim$ normal distribution for large n

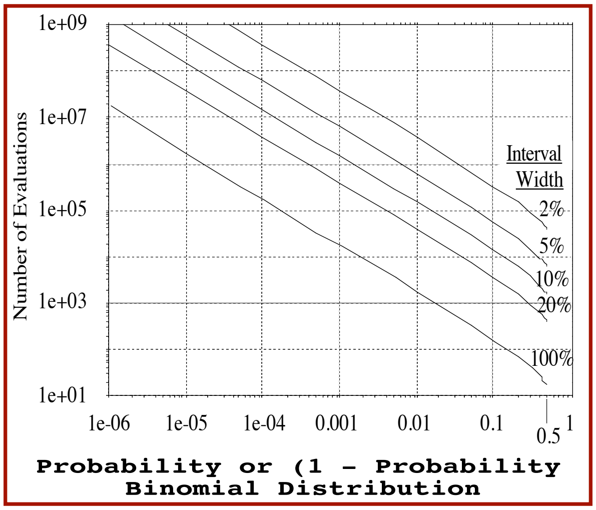


- **Design Cost Function**
  - ***High probabilities weighted more than low probabilities***
  - **$J = aPr_i^2 + bPr_{ts}^2 + c\,Pr_u^2$**
  - **$a = 1$**
  - **$b = c = 0.01$**

Monte Carlo Evaluation of Probability of Satisfying a Design Metric

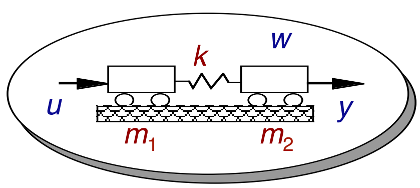
$$Pr_k(\mathbf{d},\mathbf{v}) \approx \frac{1}{N}\sum_{i=1}^{N} e_k\left[C(\mathbf{d}),H(\mathbf{v})\right], \quad k=1,3$$

$$J = aPr_i^2(\mathbf{d},\mathbf{v}) + bPr_{ts}^2(\mathbf{d},\mathbf{v}) + cPr_u^2(\mathbf{d},\mathbf{v})$$



- Compute **v** using random number generators over *N* trials
  - Required number of trials depends on outcome probability and desired confidence interval
- Search for best **d** using a genetic algorithm to minimize *J*

Uncertain Plant*



$$y = x_2 + n$$

Plant dynamic equation

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \\ \dot{x}_3 \\ \dot{x}_4 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ -k/m_1 & k/m_1 & 0 & 0 \\ k/m_2 & -k/m_2 & 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ 1/m_1 \\ 0 \end{bmatrix} u + \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1/m_2 \end{bmatrix} w$$

4th-Order Plant characteristic equation

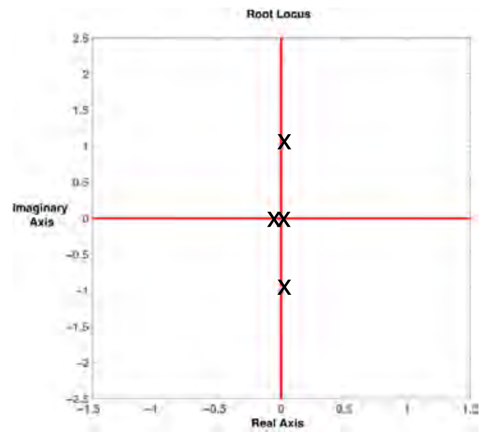
$$\Delta(s) = s^2\left[s^2 + k\frac{(m_1+m_2)}{m_1 m_2}\right] = s^2\left[s^2 - \omega_n^2\right]$$

** 1990 American Control Conference Robust Control Benchmark Problem*

Parameter Variations and Open-Loop Roots

- **Parameters of mass-spring system**
  - **Uniform probability density functions for**
    - $0.5 < m_1, m_2 < 1.5$
    - $0.5 < -k < 2$
- **Neutral stability for all mass-spring values**

$$\Delta(s) = s^2\left[s^2 + k\frac{(m_1 + m_2)}{m_1 m_2}\right]$$
$$= s^2\left[s^2 - \omega_n^{\ 2}\right]$$



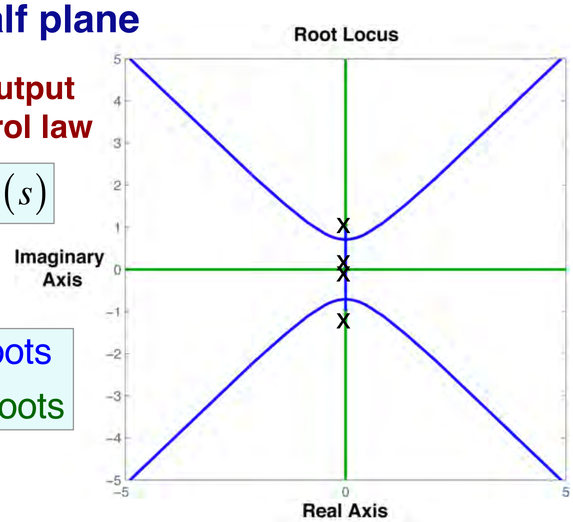
Mass-Spring-Mass Stabilization Requires Compensation

- **Proportional feedback alone cannot stabilize the system**
- **Feedback of either sign drives at least one root into the right half plane**

Single-input/single-output (SISO) feedback control law

$$u(s) = -cy(s) \triangleq -cx_2(s)$$

$c \geq 0$, Blue locus of roots
$c \leq 0$, Green locus of roots



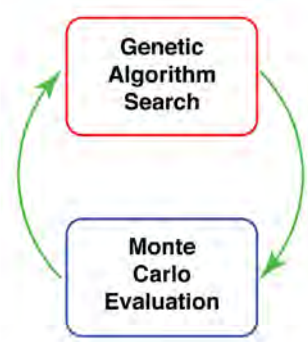
Search-and-Sweep Design of Family of Robust SISO Feedback Compensators

Begin with lowest-order feedback compensator

$$C_{12}(s) = \frac{a_0 + a_1 s}{b_0 + b_1 s + b_2 s^2} \equiv C(\mathbf{d})$$

Arrange parameters as binary design vector

$$\mathbf{d} = \left\{ a_0, a_1, b_0, b_1, b_2 \right\}$$



$$\mathbf{d}^* = \left\{ a_0^*, a_1^*, b_0^*, b_1^*, b_2^* \right\}$$

Search for design vector, *d*, that minimizes *J*

$$m_1 = rand(1) + 0.5$$
$$m_2 = rand(1) + 0.5$$
$$k = -1.5 * rand(1) + 0.5$$

Monte Carlo evaluation with uncertain parameters, v

Search-and-Sweep Design of Family of Robust Feedback Compensators

1) **Define next higher-order compensator**

$$C_{22}(s) = \frac{a_0 + a_1 s + a_2 s^2}{b_0 + b_1 s + b_2 s^2}$$

2) **Optimize over all parameters, including optimal coefficients in starting population**

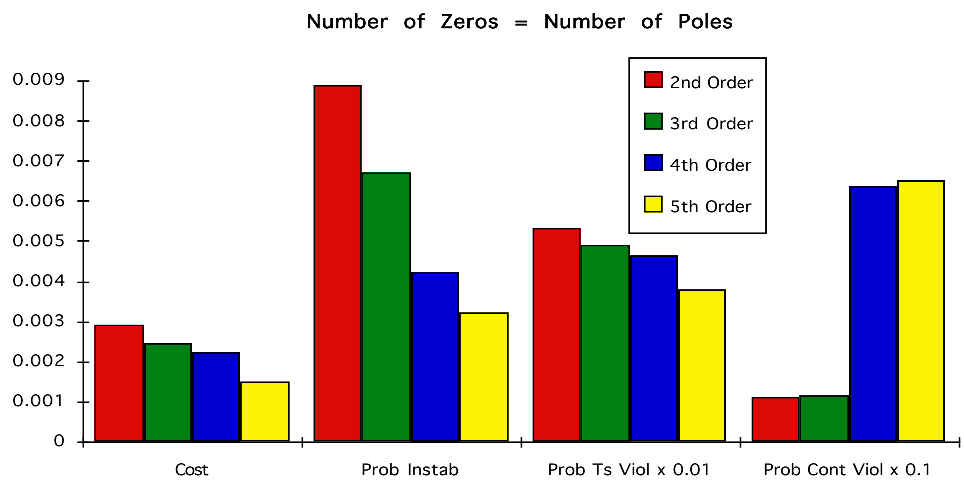
$$\mathbf{d} = \left\{ a_0^*, a_1^*, a_2, b_0^*, b_1^*, b_2^* \right\} \Rightarrow \mathbf{d}^{**} = \left\{ a_0^{**}, a_1^{**}, a_2^{**}, b_0^{**}, b_1^{**}, b_2^{**} \right\}$$

3) **Sweep to satisfactory design or no further improvement**

$$C_{23}(s) = \frac{a_0 + a_1 s + a_2 s^2}{b_0 + b_1 s + b_2 s^2 + b_3 s^3} \qquad C_{33}(s) = \frac{a_0 + a_1 s + a_2 s^2 + a_3 s^3}{b_0 + b_1 s + b_2 s^2 + b_3 s^3}$$

$$C_{34}(s) = \frac{a_0 + a_1 s + a_2 s^2 + a_3 s^3}{b_0 + b_1 s + b_2 s^2 + b_3 s^3 + b_4 s^4} \quad \ldots$$

Design Cost and Probabilities for Optimal 2nd- to 5th-Order Compensators



Next Time: Parameter Estimation and Adaptive Control

Supplemental Material

Example: Probability of Stable Control of an Unstable Plant

Longitudinal dynamics for a Forward-Swept-Wing Aircraft

$$\mathbf{F} = \begin{bmatrix} -2gf_{11}/V & \rho V^2 f_{12}/2 & \rho V f_{13} & -g \\ -45/V^2 & \rho V f_{22}/2 & 1 & 0 \\ 0 & \rho V^2 f_{32}/2 & \rho V f_{33} & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}; \quad \mathbf{G} = \begin{bmatrix} g_{11} & g_{12} \\ 0 & 0 \\ g_{31} & g_{32} \\ 0 & 0 \end{bmatrix}; \quad \mathbf{x} = \begin{bmatrix} V \\ \alpha \\ q \\ \theta \end{bmatrix}$$

Nominal eigenvalues (one unstable)

$$\lambda_{1-4} = -0.1 \pm 0.057j, \quad -5.15, \quad 3.35$$

Air density and airspeed, ρ and V, have uniform distributions(±30%)

10 coefficients have Gaussian distributions (σ = 30%)

$$\mathbf{p} = \begin{bmatrix} \rho & V & f_{11} & f_{12} & f_{13} & f_{22} & f_{32} & f_{33} & g_{11} & g_{12} & g_{31} & g_{32} \end{bmatrix}^T$$

Environment | Uncontrolled Dynamics | Control Effect

LQ Regulators for the Example

Three stabilizing 2-input feedback control laws

- **Case *a*) LQR with low control weighting**

$$\mathbf{Q} = diag(1,1,1,0); \quad \mathbf{R} = (1,1); \quad \lambda_{1-4_{nominal}} = -35, -5.1, -3.3, -.02$$

$$\mathbf{C} = \begin{bmatrix} 0.17 & 130 & 33 & 0.36 \\ 0.98 & -11 & -3 & -1.1 \end{bmatrix}$$

- **Case *b*) LQR with high control weighting**

$$\mathbf{Q} = diag(1,1,1,0); \quad \mathbf{R} = (1000,1000); \quad \lambda_{1-4_{nominal}} = -5.2, -3.4, -1.1, -.02$$

$$\mathbf{C} = \begin{bmatrix} 0.03 & 83 & 21 & -0.06 \\ 0.01 & -63 & -16 & -1.9 \end{bmatrix}$$

- **Case *c*) Case *b* with gains multiplied by 5 for bandwidth (loop-transfer) recovery**

$$\lambda_{1-4_{nominal}} = -32, -5.2, -3.4, -0.01$$

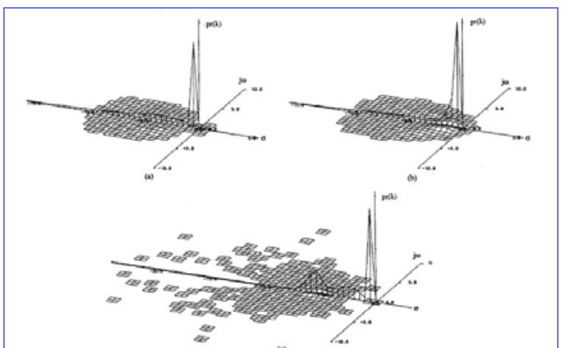
$$\mathbf{C} = \begin{bmatrix} 0.13 & 413 & 105 & -0.32 \\ 0.05 & -313 & -81 & -1.1 - 9.5 \end{bmatrix}$$

Stochastic Robustness

(Ray, Stengel, 1991)

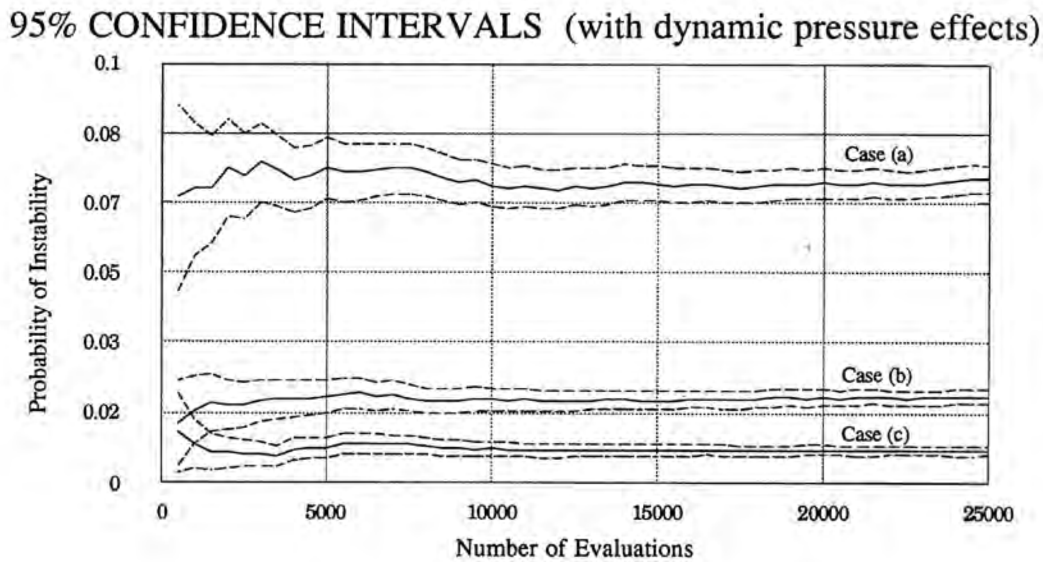
- Distribution of closed-loop roots with
  - Gaussian uncertainty in 10 parameters
  - Uniform uncertainty in velocity and air density
- 25,000 Monte Carlo evaluations

- Probability of instability
- a) Pr = 0.072
- b) Pr = 0.021
- c) Pr = 0.0076

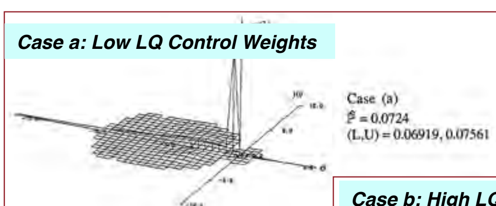


Stochastic Root Locus

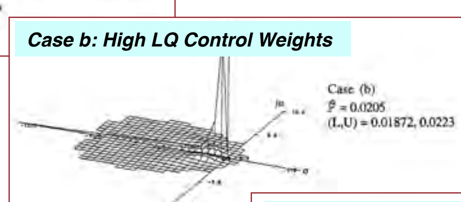
Probabilities of Instability for the Three Cases

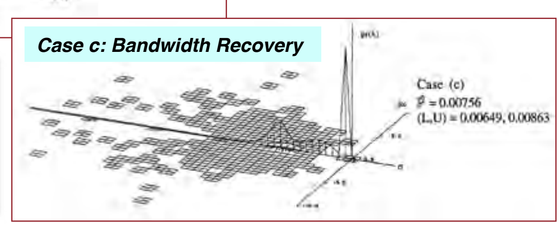


Stochastic Root Loci for the Three Cases



with Gaussian Aerodynamic Uncertainty





- Probabilities of instability with 30% uniform aerodynamic uncertainty
  - Case a: 3.4×10^{-4}
  - Case b: 0
  - Case c: 0

Markov Process

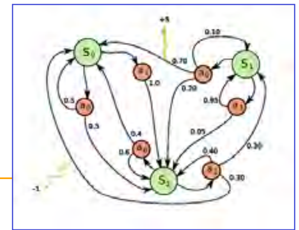
Markov Sequence and Process

- **Markov Sequence (Discrete Time)**
  - **Probability distribution of dynamic process at time $t_{k+1} > t_k > 0$, conditioned on the past history**
  - **Depends only on the state, x, at time t_k**

$$\Pr\left[x_{k+1} \mid \left(x_k, x_{k-1}, x_{k-2}, \cdots, 0\right)\right] = \Pr\left[x_{k+1} \mid x_k\right]$$

- **Markov Process (Continuous Time)**
  - **Probability distribution of dynamic process at time $s > t > 0$, conditioned on the past history**
  - **Depends only on the state, x, at time t**

Markov Decision Sequence



- **Model for decision making under uncertainty contains following elements**

$$\left[\mathbf{X},\mathbf{A},P_{a_m}(\mathbf{x}_i,\mathbf{x}'),L_{a_m}(\mathbf{x}_i,\mathbf{x}')\right]$$

where

$\mathbf{X}$: **Finite set of states**, $\mathbf{x}_1,\ \mathbf{x}_2,\ldots\mathbf{x}_i,\ldots,\mathbf{x}_I$

$\mathbf{A}$: **Finite set of actions**, $\mathbf{a}_1,\mathbf{a}_2,\ldots,\mathbf{a}_j,\ldots,\mathbf{a}_J$

$$P_{\mathbf{a}_j}(\mathbf{x}_k,\mathbf{x}') = \Pr\left\{\left[\mathbf{x}(t_{k+1})=\mathbf{x}'\right]\left[\mathbf{x}(t_k)=\mathbf{x}_k \text{ and } \mathbf{a}(t_k)=\mathbf{a}_j\right]\right\}$$

$= \text{Probability that } \mathbf{a}_j \text{ will cause } \mathbf{x}_i(t_k) \text{ to transition to } \mathbf{x}'$

$L_{\mathbf{a}_j}(\mathbf{x}_k,\mathbf{x}') =$ **Expected immediate reward** for transition from $\mathbf{x}_k$ to $\mathbf{x}'$

- **Optimal decision maximizes (minimizes) expected total reward (cost) by choosing best set of actions (control policy)**
  - Linear-quadratic-Gaussian (LQG) control
  - Dynamic programming -> HJB equation ~> A* search
  - Reinforcement learning ~> Heuristic search

Maximizing the Utility Function of a Markov Process

Utility function: $$J = \lim_{k_f\to\infty}\sum_{k=0}^{k_f}\gamma(t_k)L_{\mathbf{a}}\left[\mathbf{x}(t_k),\mathbf{x}(t_{k+1})\right]$$

$\gamma(t_k)$: **Discount rate**, $0<\gamma(t_k)<1$

Utility function to go = Value function:

$$V = \lim_{k_f\to\infty}\sum_{k=k_{current}}^{k_f}\gamma(t_k)L_{\mathbf{a}}\left[\mathbf{x}(t_k),\mathbf{x}(t_{k+1})\right]$$

Maximizing the Utility Function of a Markov Process

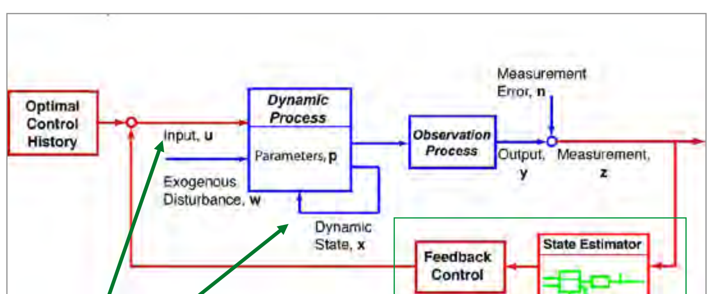
Optimal control at *t*

$$\mathbf{u}_{opt}\left(t_k\right) = \arg\max_{\mathbf{a}} \left\{ L_{\mathbf{a}}\left[\mathbf{x}(t_k),\mathbf{x}(t_{k+1})\right] + \gamma(t_k) \sum_{k=k_{current}}^{\infty} P_{\mathbf{a}}\left[\mathbf{x}(t_k),\mathbf{x}(t_{k+1}))\right] V\left[\mathbf{x}(t_{k+1})\right] \right\}$$

Optimized value function

$$V^*\left(t_k\right) = L_{\mathbf{u}_{opt}(t_k)}\left[\mathbf{x}^*(t_k)\right] + \gamma(t_k) \sum_{k=k_{current}}^{\infty} P_{\mathbf{u}_{opt}(t_k)}\left[\mathbf{x}^*(t_k),\mathbf{x}_{est}^*(t_{k+1})\right] V\left[\mathbf{x}^*_{est}(t_{k+1})\right]$$

LQG Control Optimizes Discrete-Time LTI Markov Process



LQG Gain Calculation

$$\left[\mathbf{X},\mathbf{A},P_{a_m}\left(\mathbf{x}_i,\mathbf{x}'\right),L_{a_m}\left(\mathbf{x}_i,\mathbf{x}'\right)\right]$$

where

$\mathbf{X}$: **Finite set of states**, $\mathbf{x}_1, \mathbf{x}_2,\ldots\mathbf{x}_i,\ldots,\mathbf{x}_I$

$\mathbf{A}$: **Finite set of actions**, $\mathbf{a}_1,\mathbf{a}_2,\ldots,\mathbf{a}_j,\ldots,\mathbf{a}_J$

$P_{\mathbf{a}_j}\left(\mathbf{x}_k,\mathbf{x}'\right) = \Pr\left\{\left[\mathbf{x}\left(t_{k+1}\right)=\mathbf{x}'\right] \middle| \left[\mathbf{x}\left(t_k\right)=\mathbf{x}_k \text{ and } \mathbf{a}\left(t_k\right)=\mathbf{a}_j\right]\right\}$

= Probability that $\mathbf{a}_j$ will cause $\mathbf{x}_i\left(t_k\right)$ to transition to $\mathbf{x}'$

$L_{\mathbf{a}_j}\left(\mathbf{x}_k,\mathbf{x}'\right) =$ **Expected immediate reward** for transition from $\mathbf{x}_k$ to $\mathbf{x}'$