Lectures on Stochastic System Analysis and Bayesian Updating

June 29-July 13 2005

James L. Beck, California Institute of Technology
Jianye Ching, National Taiwan University of Science & Technology
Siu-Kui (Ivan) Au, Nanyang Technological University, Singapore

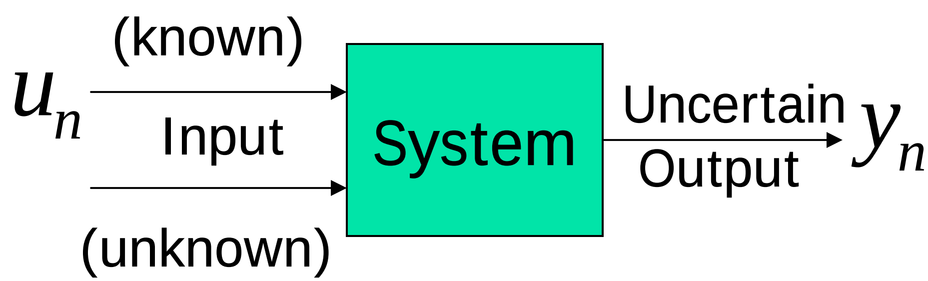
Overview of Robust Stochastic System Analysis

- **1.1: Stochastic System Modeling**
  - Class(es) of probabilistic input-output models for system to address uncertainties in system modeling (robust analysis)
  - Terminology: ‘System’= real thing; ‘Model’=idealized mathematical model of system
- **1.2: Prior System Analysis**
  - Uncertainties in system input also addressed
  - Reliability analysis to compute failure probabilities
- **1.3: Posterior System Analysis**
  - Bayesian updating of models in class based on system data
  - Updated reliability analysis

1.1: Stochastic System Modeling

- **Predictive model:** Gives probabilistic input-output relation for system depending on model parameters:

$$p(Y_n \mid U_n, \theta)$$



where input (if available) and output time histories:

$$U_n = \{u_k \in R^{N_I} : k = 0, \ldots, n\}$$

$$Y_n = \{y_k \in R^{N_O} : k = 1, \ldots, n\}$$

Stochastic System Modeling (Continued)

- Usually have a set of possible predictive probability models to represent system:

$$\{\, p(Y_n \mid U_n, \theta) : \theta \in \Theta \subset R^{N_p} \,\}$$

  - **Nominal prior predictive model**: Select single model, e.g. most plausible model in set
  - But there is uncertainty in which model gives most accurate predictions that should not be ignored

Stochastic System Modeling (Continued)

- **Robust prior predictive model**:

  Select $p(\theta \mid \mathrm{M})$ to quantify the plausibility of each model in set, then from Total Probability Theorem:

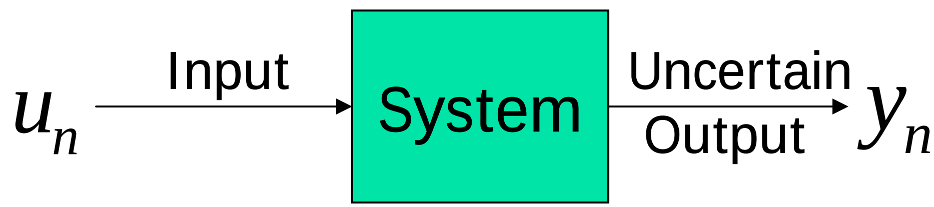
$$p(Y_n \mid U_n, \mathrm{M}) = \int p(Y_n \mid U_n, \theta)\, p(\theta \mid \mathrm{M})\, d\theta$$

Here, M denotes the class of probability models, i.e. it specifies the functional forms of $p(Y_n \mid U_n, \theta)$ & $p(\theta \mid \mathrm{M})$

More about choosing $p(\theta \mid \mathrm{M})$, the prior PDF, later

Stochastic System Model: Example 1

- **Complete system input known:** Define deterministic input-output model $q_n(U_n,\theta)$ for $\theta \in \Theta \subset R^{N_p}$

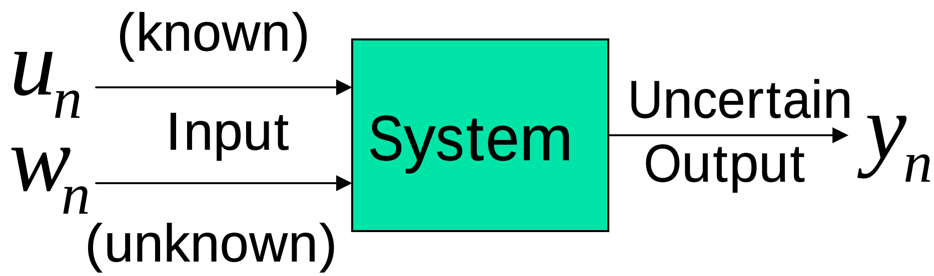


Uncertain prediction error: $v_n = y_n - q_n(U_n,\theta)$

Model for prediction-error time history gives $p(Y_n \mid U_n,\theta)$
Can take prediction errors as zero-mean Gaussian & independent in time (maximum entropy distribution), so Y_n is Gaussian with mean $q_n(U_n,\theta)$ and covariance matrix $\Sigma(\theta)$

Stochastic System Model: Example 2

- **Complete system input not known:** Define state-space dynamic model for system by:

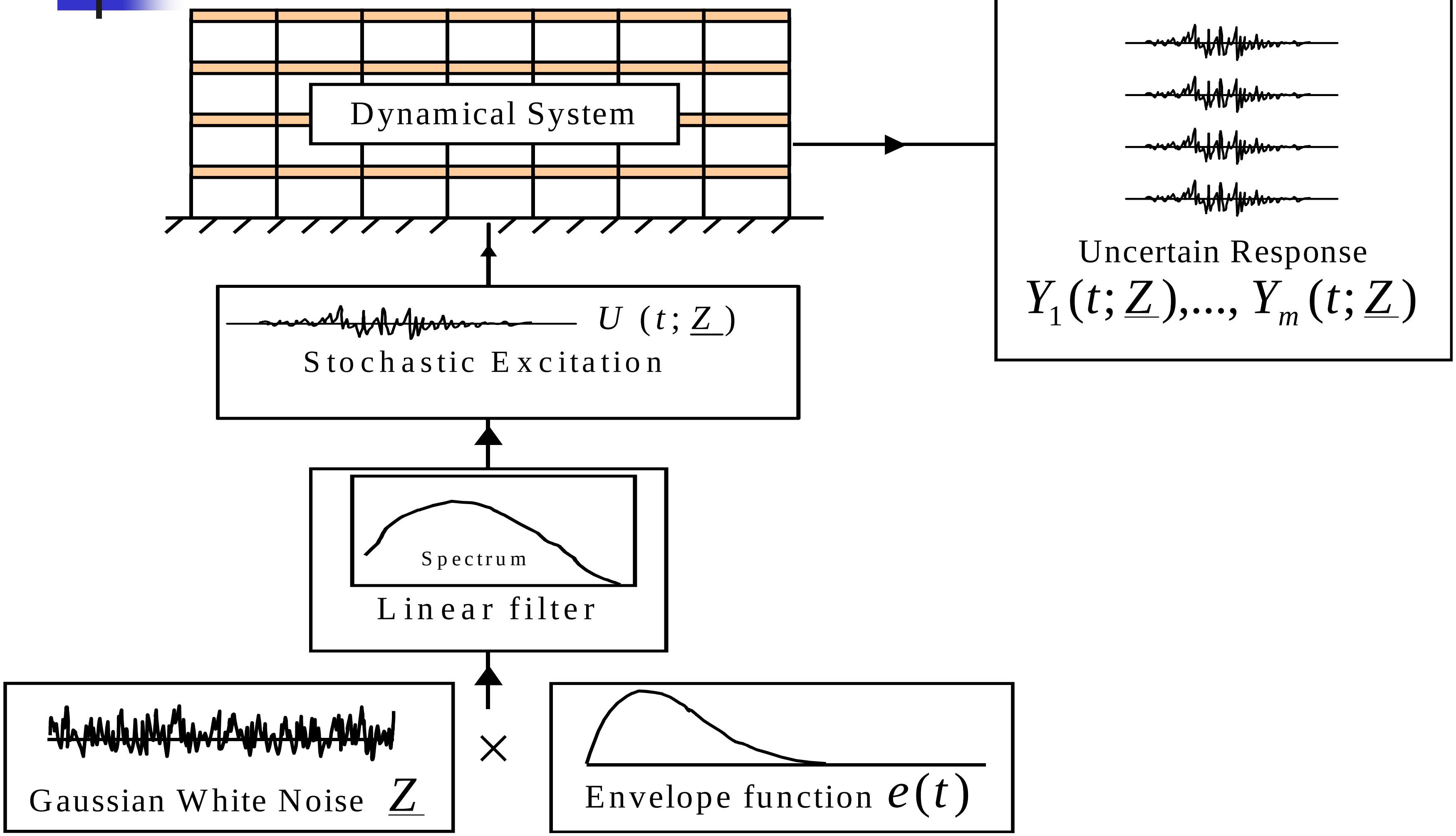


Uncertain state: $$x_n = F(x_{n-1}, u_{n-1}, w_{n-1}, \theta)$$

Uncertain output: $$y_n = H(x_n, u_n, \theta) + v_n$$

Probability models for missing information (i.e. initial state x_0 and time histories of unknown input w_n and prediction error v_n), define $p(Y_n \mid U_n, \theta)$

First-Excursion Problem: Analysis Model



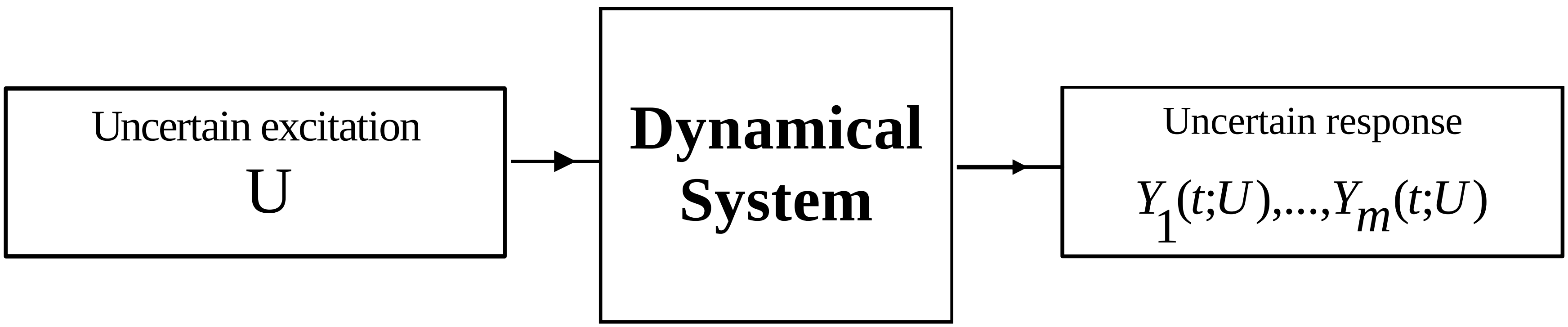
1.2: Prior System Analysis

- **Total Input Uncertainty:** Choose probability model over set of possible system inputs: $p(U_n \mid \mathrm{U})$
- **Nominal prior predictive analysis**: Find the probability that system output lies in specified set F using nominal model:

$$P(Y_n \in F \mid \mathrm{U}, \theta) = \int P(Y_n \in F \mid U_n, \theta)\, p(U_n \mid \mathrm{U})\, dU_n$$

  - Reliability problem corresponds to $Y_n \in F$ defining 'failure' (= specified unacceptable performance of system)
  - Primary computational tools for complex dynamical systems are advanced stochastic simulation methods (more later) and Rice's out-crossing theory for simpler systems

First-Excursion Problem



$$P(\text{Failure}) = P\left(\bigcup_{i=1}^{m}\{|Y_i(t;U)| > b_i \text{ for some } t \text{ on } [0,T]\}\right)$$

Prior System Analysis (Continued)

- **Robust prior predictive analysis**:

$$P(Y_n \in F \mid \mathrm{U}, \mathrm{M}) = \int P(Y_n \in F \mid \mathrm{U}, \theta)\, p(\theta \mid \mathrm{M})\, d\theta$$

  - Robust reliability if $Y_n \in F$ defines failure
  - Primary computational tools:
    - Stochastic simulation, e.g. importance sampling with ISD at peak(s) of integrand (needs optimization)
    - Asymptotic approximation w.r.t. curvature of the peak(s) of integrand (needs optimization)
  - Huge differences possible between nominal and robust failure probabilities

Prior System Analysis (Continued)

- Asymptotic approximation introduced in:
  - Papadimitriou, Beck and Katafygiotis (1997). “Asymptotic expansions for reliability and moments of uncertain systems.” (at website)
  - Au, Papadimitriou and Beck (1999). “Reliability of Uncertain Dynamical Systems with Multiple Design Points” (at website)
- Comparisons between nominal and robust failure probabilities available in:
  - Papadimitriou, Beck & Katafygiotis (2001). “Updating Robust Reliability using Structural Test Data.” (at website)

1.3: Posterior System Analysis

- **Available System Data:** $\mathrm{D}_{\mathrm{N}} = \{U_{\mathrm{N}}, Y_{\mathrm{N}}\}$
- **Update by Bayes Theorem:**

$$p(\theta \mid \mathrm{D}_{\mathrm{N}}, \mathrm{M}) = c\,p(Y_{\mathrm{N}} \mid U_{\mathrm{N}}, \theta)\, p(\theta \mid \mathrm{M})$$

- **Optimal posterior predictive model**:

  Select most plausible model in class based on data,
  i.e. $\hat{\theta}$ that maximizes the posterior PDF (if unique)

- **Optimal posterior predictive analysis:**

$$P(Y_n \in F \mid \mathrm{U}, \hat{\theta}) = \int P(Y_n \in F \mid U_n, \hat{\theta})\, p(U_n \mid \mathrm{U}\ )\, dU_n$$

  - Difficulties: Non-convex multi-dimensional optimization (‘parameter estimation’); ignores model uncertainty

Posterior System Analysis (Continued)

- **Robust posterior predictive model:**

  Use all predictive models in class weighted by their updated probability (exact solution based on probability axioms):

$$p(Y_n \mid U_n, \mathrm{D_N}, \mathrm{M}) = \int p(Y_n \mid U_n, \theta)\, p(\theta \mid \mathrm{D_N}, \mathrm{M})\, d\theta$$

- **Robust posterior predictive analysis**:

$$P(Y_n \in F \mid \mathrm{U}, \mathrm{D_N}, \mathrm{M}) = \int P(Y_n \in F \mid \mathrm{U}, \theta)\, p(\theta \mid \mathrm{D_N}, \mathrm{M})\, d\theta$$

  - Primary computational tools are MCMC simulation methods and asymptotic approximation w.r.t. sample size N

Posterior System Analysis (Continued)

- **Asymptotic approximation for large N for robust posterior predictive analysis**

  (Beck & Katafygiotis 1998; Papadimitriou, Beck & Katafygiotis 2001 - both at website)

$$P(Y_n \in F \mid \mathrm{U}, \mathrm{D}_{\mathrm{N}}, \mathrm{M}) = \int P(Y_n \in F \mid \mathrm{U}, \theta)\, p(\theta \mid \mathrm{D}_{\mathrm{N}}, \mathrm{M})\, d\theta$$

$$\approx \sum_{k=1}^{K} w_k P(Y_n \in F \mid \mathrm{U}, \hat{\theta}_k)$$

  - Assumes system is identifiable based on the data, i.e. finite number of MPVs $\hat{\theta}_1, \hat{\theta}_2, \ldots, \hat{\theta}_K$ that locally maximize posterior PDF, so need to do optimization; uses Laplace's method for asymptotic approximation (see later)

Posterior System Analysis (Continued)

- The weights w_k are proportional to the volume under the peak of the posterior PDF at $\hat{\theta}_k$ and sum to unity (see Beck & Katafygiotis 1998)
- Globally identifiable case (K=1) justifies using MPV $\hat{\theta}$ for posterior predictive model when there is large amounts of data:

$$p(Y_n \,|\, U_n, \mathrm{D_N}, \mathrm{M}) \approx p(Y_n \,|\, U_n, \hat{\theta})$$

  - Gives a rigorous justification for doing predictions with MPV model (or MLE, since $\hat{\theta}$ is insensitive to choice of prior)
  - Error in approximation is $O(1/N)$

Posterior System Analysis (Continued)

- Unidentifiable case corresponds to a continuum of MPVs lying on a lower dimensional manifold in the parameter space
  - Interest in this case is driven by finite-element model updating
  - Asymptotic approximation for posterior predictive model for large amount of data is an integral over this manifold – feasible if it is low dimension (<4?) (Katafygiotis and Lam (2002); Papadimitriou, Beck and Katafygiotis (2001) - both at website)
  - All MPV models give similar predictions at observed DOFs but may be quite different at unobserved DOFs

Posterior System Analysis (Continued)

- **Stochastic Simulation approaches:**
  - Very challenging because most of probability content of posterior PDF is concentrated in a small volume of parameter space (IS does not work)
  - But potential of avoiding difficult non-convex multi-dimensional optimization and handling unidentifiable case in higher dimensions
  - Markov Chain Monte Carlo simulation (e.g. Metropolis-Hastings algorithm) shows promise (more later)

Comments

- The framework and computational tools give a powerful approach to stochastic system analysis and yet it is not widely used in engineering – why not?
- Obstacle: many people are comfortable with $p(Y_n | U_n, \theta)$ but not with $p(\theta | \mathrm{M})$ because they interpret probability as the relative frequency of inherently random events in the long run

Comments

- The two main themes for the remaining lectures:
  - Development of probability logic which gives a rigorous framework in which probabilities of models makes sense
  - Development of a set of computational tools to provide efficient algorithms for handling the high-dimensional algorithms needed for prior and posterior stochastic predictive system analysis

Probability Logic

- Primarily due to:

  R.T. Cox 1946, 1961: The Algebra of Probable Inference

  E.T. Jaynes 1983, 2003: Probability Theory – The Logic of Science

- Major contributors to development of ideas:

  T. Bayes 1763: An essay towards solving a problem in the doctrine of chances

  P.S. Laplace 1812: Analytical Theory of Probability

  H. Jeffreys 1931: Scientific Inference

  1939: Theory of Probability

- Quote from James Clerk Maxwell (1850):

  The actual science of logic is conversant at present only with things either certain, impossible or entirely doubtful, none of which (fortunately) we have to reason on. Therefore the true logic of this world is the calculus of probabilities, which takes account of the magnitude of the probability which is, or ought to be, in a reasonable man's mind.

Introduction

- **Features of Probability Logic**
  - Probability logic is a quantitative approach to plausible reasoning when available information is incomplete; it generalizes binary Boolean logic
  - Framework based on probability axioms and no other ad-hoc criteria or concepts
  - Uses Cox-Jaynes interpretation of probability as quantifying plausibility of statements conditional on specified information
  - Probability models are used to stand in for missing information; they are (lack of) knowledge models

Introduction (Continued)

- **Features of Probability Logic**
  - Careful tracking of all conditioning information since all probabilities are conditional on probability models and other specified information
  - Meaningful to talk about probability of probability models, an essential aspect of Bayesian analysis
  - Involves integrations over high-dimensional input and model parameter spaces; computational tools for this will be given and are also being actively developed by many researchers
  - Framework is general but our focus is primarily on dynamical systems

Decision Making under Uncertainty/Incomplete Information (e.g. Engineering System Design)

System Models and Probability Models
[Knowledge models for missing information]

→

Information Processing
[Axioms for Calculus of Probability]

→

Decision Making
[Predicted Decision Variables]

Information processing should be done in such a way that known information is not lost and spurious information is not added