# CONTROL OF STOCHASTIC SYSTEMS

**P.R. Kumar**
*Department of Electrical and Computer Engineering, and Coordinated Science Laboratory, University of Illinois, Urbana-Champaign, USA.*

**Keywords:** Markov chains, transition probabilities, controlled Markov chains, noisy observations, partially observed systems, linear stochastic systems, linear Gaussian systems, controlled auto regressive models, autoregressive moving average systems with exogenous inputs, cost functions, time-horizon, terminal cost, running cost, state-feedback policy, Markov policy, optimal cost-to-go, dynamic programming equation, principle of optimality, linear quadratic Gaussian problem, equilibrium points, stability, countable Markov chains, steady-state, super-martingales, Lyapunov functions, stopping times, recurrence, positive recurrence, stochastic stability, estimation, Bayes Rules, Kalman filter, minimum mean-square error estimate, conditional mean, conditional covariance, posterior probability distribution, Bayesian and Non-Bayesian approaches to adaptive control, parameter vector, maximum likelihood, least squares estimate, prediction error estimate, system identification, consistency, separation theorem, certainty equivalence, information state, hyperstate, self-tuning regulators.

## Contents





## Summary

We present an account of several topics, modeling, control, estimation, stability, identification and adaptive control, which arise in the study of the control of stochastic systems.

## 1. Introduction

A holistic treatment of the problem of control of stochastic systems encompasses the following topics:

(i) Models of stochastic systems

(ii) Optimal stochastic control

(iii) Stability of stochastic systems

(iv) Estimation of stochastic systems

(v) Identification of stochastic systems

(vi) Control of partially observed systems

(vii) Adaptive control

We present an outline of each of these topics which will enable the reader to obtain an integrated perspective of the field.

**2. Models of Stochastic Systems**

A discrete-time stochastic process $\{x(t)\}_{t=0}^{+\infty}$ is a *Markov chain* if $p(x(t+1)|x(0),\ldots,x(t)) = p(x(t+1)|x(t))$. That is, the conditional distribution of the future state $x(t+1)$ depends on the past $(x(0),\ldots,x(t))$ only through the present state $x(t)$.

Indeed this justifies the use of the name "state". The Markov chain can then be described by its *transition probabilities* $p(x(t+1)|x(t))$.

Extending this notion, one can describe a *controlled* Markov chain by its *controlled* transition probabilities $p(x'|x, u)$ which describe the conditional probability of the next state $x(t+1)$ being $x'$, when the current state $x(t)=x$, and an input $u(t)=u$ is applied.

If the state $x(t)$ is not observed, then it is common to model the observations $y(t)$ by the conditional probability distribution $p(y|x)$ which describes the probability distribution of the observation $y(t)$ when the state $x(t)=x$.

The system is then called a *partially observed controlled Markov chain.*

If the transition probabilities depend on the time $t$, then one can describe the time-varying system by the pair of transition probabilities $p(x'|x, u, t)$ and $p(y|x,t)$.

A common deterministic noise-free state space model of a systems in discrete-time is

$$\begin{aligned} x(t+1) &= f(x(t), u(t), t) \\ y(t) &= g(x(t), t), \end{aligned}$$

where $x(t)$ is the state of the system at time $t$, $u(t)$ is the input applied at time $t$, and $y(t)$ is the output at time $t$.

The corresponding stochastic analog of the *state space model* is

$$\begin{aligned} x(t+1) &= f(x(t),\ u(t),\ w(t),\ t) \\ y(t) &= g(x(t), v(t)\ t), \end{aligned}$$

where $w(t)$ is the noise entering the state equation, and $v(t)$ is the noise entering the observation equation. These noises are modeled as stochastic processes (see *Models of Stochastic Systems*).

If $\{w(0), w(1), w(2), \ldots\}$ are mutually independent, then $x(t)$ is indeed the state of a controlled Markov chain. If further $\{v(0),\ v(1), v(2), \ldots\}$ are also mutually independent, and $w, v$ are independent of each other, then one has a partially observed controlled Markov chain written in the form of state and observation equations.

If $\{w(0), w(1), w(2), \ldots\}$ are not mutually independent, then one often models them as the output of a system driven by independent random variables $\{n(0), n(1), \ldots, m(0), m(1), \ldots\}$

$$\begin{aligned} z(t+1) &= h(z(t),\ n(t)) \\ w(t) &= k(z(t), m(t)). \end{aligned}$$

In such situations, one can adjoin $z$ to $x$ and let $(x, z)$ serve as the state.

A special and important case of such a state space model is a *linear stochastic system*:

$$\begin{aligned} x(t+1) &= A(t)\,x(t) + B(t)u(t) + G(t)w(t) \\ y(t) &= C(t)\,x(t) + H(t)v(t), \end{aligned}$$

where $A(t), B(t), C(t), G(t)$ and $H(t)$ are time-varying matrices of appropriate dimensions. A model that particularly lends itself to analysis is when the noise processes $w(t)$ and $v(t)$ are jointly Gaussian stochastic processes (see *Models of Stochastic Systems*). Then it is called a *Linear-Gaussian* model.

Instead of dealing with the state $x(t)$, one can directly model how the input influences the output, i.e., by an *input-output* model. The most common model is a Control Autoregressive Moving Average Model (CARMA) or Autoregressive Moving Average Model with Exogenous Inputs (ARMAX) model:

$$\begin{aligned} y(t) + a_1 y(t-1) + \cdots + a_n y(t-n) = b_0 u(t) + b_1 u(t-1) + \cdots \\ + b_n u(t-n) + w(t) + c_1 w(t-1) + \cdots + c_n w(t-n). \end{aligned}$$

One can also consider the continuous time counterpart of the state-space model:

$$\begin{aligned} dx(t) &= f(x(t),\ u(t),\ t)dt + \sigma(x(t))\,dw(t) \\ dy(t) &= g(x(t))\,dt + dv(t). \end{aligned}$$

Here $w(t)$ and $v(t)$ are Brownian motion processes, and one has to interpret the above *stochastic differential equations* in the appropriate mathematical way. This requires a knowledge of Ito stochastic integrals and stochastic calculus.

## 3. Optimal Stochastic Control

Consider the case of a discrete-time stochastic system where the state $x(t)$ is directly observed. How should one choose the control input $\{u(t)\}$ to be applied to such a system? A common approach is to consider a *cost function* of the form

$$E\left[h(x(T+1))+\sum_{t=0}^{T}c(x(t),\, u(t))\right],$$

and choose control inputs which minimize this expected cost. Above $T$ is a *time horizon*, $h(x(T+1))$ is the *terminal cost*, and $c(x(t), u(t))$ is the *running cost*. One minimizes this cost over the set of history dependent *strategies* where $u(t)=u(x(0),\ldots,x(t),t)$ is allowed to depend on the entire past of the observations and the current time $t$. It can be shown that within this class of history dependent strategies one can restrict attention to strategies of the form $u(t)=u(x(t),t)$ where the input depends only on the current state and current time. Such a strategy can be termed as a *state feedback* policy or a *Markov* policy.

If one defines the *optimal remaining cost* or *optimal cost-to-go* from a state $x$ at time $t$ by

$$V(x,t) := \mathrm{Min}_{u(\cdot)}E\left[h(x(T+1))+\sum_{s=t}^{T}c(x(s),\, u(s))\,\middle|\,x(t)=x\right],$$

then it can be shown that this function satisfies the following equation:

$$V(x,t)=\mathrm{Min}_{u}\left\{c(x,u)+\sum_{x'}p(x'\,|\,x,\,u)V(x',\,t+1)\right\},$$

with the terminal condition

$$V(x,\,T+1)=h(x).$$

Essentially the above equation says that the optimal cost from a state $x$ at time $t$ is obtained by considering different choices of an input $u$ to apply at time $t$. For each such potential input $u$, one determines the current cost $c(x,u)$ as well as the *expected* cost from the state reached at the next time instant. Then, one simply chooses the best input to apply at the present time as the one which minimizes the sum of the expected current cost plus the expected remaining cost. This equation is called the *dynamic programming equation*, and the logic leading to it as the *principle of optimality*. It also follows that if for $(x,t)$ one chooses the minimizing $u$, calling it $u(x,t)$, then $u(x,t)$ is the optimal policy. Thus the optimal policy can be chosen as a Markov or a state feedback policy.

The dynamic programming approach can be extended to other models and situations, as shown in *Dynamic Programming*.

A particular special case of great interest in control is the so-called *Linear-Quadratic-Gaussian (LQG) problem*. For a linear system with independent white Gaussian noises $w$ and $v$,

$$\begin{aligned} x(t+1) &= A(t)x(t) + B(t)u(t) + G(t)w(t) \\ y(t) &= C(t)x(t) + H(t)v(t), \end{aligned}$$

one seeks to minimize a *quadratic* cost criterion:

$$E\left[ x^T(T+1)Sx(T+1) + \sum_{t=1}^{T} (x^T(t)Q(t)x^T(t) + u^T(t)R(t)u(t)) \right],$$

where $S \geq 0,\ Q(t) \geq 0$ and $R(t) > 0.$

The cost-to-go function turns out to be quadratic function of the state plus a deterministic term:

$$V(x,t) = x^T S(t)x + \gamma(t).$$

By substituting this form in the dynamic programming equation, one can solve for $S(t)$ and $\gamma(t)$ in terms of $S(t+1)$ and $\gamma(t+1)$ (remember that dynamic programming solves the problem backwards in time). With the boundary conditions $S(T+1) = S$ and $\gamma(T+1) = 0,$ one thus obtains recursions for $S(t)$ and $\gamma(t)$. From the minimizing argument in the dynamic programming equation one also determines that the optimal control law is of the form

$$u(t) = K(t)x(t),$$

i.e., linear time varying feedback, with $K(t)$ expressible in terms of $S(t)$. The LQG problem thus admits a clean solution. The details of the solution are given in *LQ-stochastic Control*. Given that the quadratic cost function is a reasonable criterion, and given the widespread usage of linear models, this solution has proved to be eminently useful in control system design.

In many situations of interest, e.g., in adaptive control and self-tuning regulators, see *Self-Tuning Control*, one wishes to work with input-output models with quadratic costs. This is dealt with in *Minimum Variance Control*.

-
-
-



## Bibliography

B. Anderson and J.B. Moore, *Optimal Filtering*. Englewood Cliffs, NJ: Prentice-Hall, 1979. [Contains a comprehensive treatment of estimation for linear stochastic systems].

D.P. Bertsekas, *Dynamic Programming*: *Deterministic and Stochastic Models*. Englewood Cliffs, NJ: Prentice-Hall, 1987. [A comprehensive treatment of discrete-time dynamic programming].

G.C. Goodwin and K.S. Sin, *Adaptive Filtering, Prediction and control*. Englewood Cliffs, NJ: Prentice-Hall, 1984. [Contains a treatment of discrete-time modeling of linear systems, as well as identification and adaptive control].

P.R. Kumar and P.P. Varaiya, *Stochastic Systems: Estimation, Identification and Adaptive Control*. Englewood Cliffs, NJ: Prentice-Hall, 1986. [Contains a concise treatment of several topics including modeling, control, estimation, identification, and adaptation].

H. Kushner, *Introduction to Stochastic Control*. Hold, Rinehart and Winston, 1971. [Contains a treatment of stochastic control as well as stochastic stability].

L.Ljung and T. Söderstrom, *Theory and Practice of Recursive Identification*. Cambridge, MA: MIT Press, 1983. [Contains a treatment of identification and parameter estimation for linear systems].

## Biographical Sketch

**P.R. Kumar** is the Franklin W. Woeltge Professor of Electrical and Computer Engineering, and a Research Professor in the Coordinated Science Laboratory, at the University of Illinois, Urbana-Champaign. He was the recipient of the Donald P. Eckman Award of the American Automatic Control Council. He has presented plenary lectures at the SIAM Annual Meeting and the SIAM Control Conference in 2001, the IEEE Conference on Decision and Control in San Antonio, Texas, 1993, the SIAM Conference on Optimization in Chicago, 1992, the SIAM Annual Meeting at San Diego, 1994, Brazilian Automatic Control Congress, and the Third Annual Semiconductor Manufacturing, Control and Optimization Workshop. He is co-author with Pravin Varaiya of the book, “Stochastic Systems: Estimation, Identification and Adaptive Control”. He serves on the editorial boards of Communications in Information and Systems, Journal of Discrete Event Dynamic Systems; Mathematics of Control Signals and Systems; Mathematical Problems in Engineering: Problems, Theories and Applications; and in the past has served as Associate Editor at Large of IEEE Transactions on Automatic Control; Associate Editor of SIAM Journal on Control and Optimization; Systems and Control Letters; Journal of Adaptive Control and Signal Processing; and the IEEE Transactions on Automatic Control. He is a Fellow of IEEE. Professor Kumar’s current research interests are in wireless networks, distributed real-time systems wafer fabrication plants, and machine learning.