STOCHASTIC STABILITY

H.J. Kushner
Applied Mathematics, Brown University, Providence, RI, USA.

Keywords: stability, stochastic stability, random perturbations, Markov systems, robustness, perturbed systems, Liapunov functions, stochastic Liapunov functions, perturbed Liapunov functions, local contraction, invariant set, stochastic invariant set, martingales, supermartingales, convergence with probability one, asymptotic convergence with probability one, exponential convergence, stability of the origin, recurrence, moment boundedness, randomly sampled systems, recursive algorithms, parameter identification.

Contents

Summary

The article surveys the basic theory of stochastic stability of dynamical systems via the stochastic Liapunov function approach. The ideas provide a powerful methodology for determining the long-term properties under broad conditions, whether actual convergence, insensitivity to perturbations, boundedness or recurrence. Such properties are fundamental attributes of well-designed systems, adaptive algorithms, etc. The article starts with a quick review of the deterministic theory from the point of view that will be needed to understand and guide the basic elements of the stochastic theory. Then, the basic results for the stochastic theory are given. This is done for discrete time processes to avoid the technicalities of the continuous time theory. However, the essential ideas and results are similar. The standard types of convergence and recurrence are covered as is the stochastic invariant set theorem. The ideas are illustrated with examples drawn from randomly sampled systems, recursive optimization and identification theory. The examples serve as a useful vehicle for illustrating the diverse applications, as well as for the introduction of the perturbed Liapunov function method, which allows treatment of correlated noise.

1. Introduction: The Stochastic Stability Problem

Stability concerns the long-term, qualitative properties of dynamical physical systems. Typical questions are whether the paths converge to some limit point or set, or whether the paths are bounded or unbounded or are insensitive to perturbations. In the first case, we might wish to characterize the limit sets and we might want to know some measure of the rate of convergence. In the second case, one wishes to have good bounds or useful characterizations of the paths arbitrarily far into the future, or estimates on the rate of divergence, if appropriate. Stability is, perhaps, the quintessential problem of control theory; since the main questions traditionally have been concerned with potentially destabilizing effects of controls or environmental factors. Stability is often the first step to showing that a process is well defined and can be studied over a long time interval.

Stability issues arise in numerous applications. In adaptive control theory, one is concerned with convergence of the adaptive algorithm; and, this is generally a problem in stability, where convergence is measured in terms of an error function. Indeed, stability is one of the primary concerns throughout adaptation and learning theory. Once stability is shown (for example, that the system errors are bounded), the convergence proofs are much simpler. These learning problems are usually stochastic, since the data is, and one needs effective methods, which yield useful results in a stochastic environment.

Control systems are subject to random time variations or random errors of many types, whether in the system itself or in the physical environment in which it operates. The system might be subject to random sampling, parameters might drift or age randomly with time, or there might be random variations in the forces affecting it, etc. Stability issues arise in economics and biological and environmental modeling as well. Many economic models are based on adjustment and response due to repeated interaction, such as between, buyer and seller. In such cases, the long term behavior (such as, convergence, or whether the prices or supplies stay within certain bounds, explode, or converge to some limit) is of interest; particularly, when the adjustments in the economic activity of concern is based on random data or imperfect information. Further applications occur in queueing and in other areas of operations research.

This article will outline typical aspects of the theory and results, from an elementary viewpoint, using mostly discrete time systems. The probability theory, which is involved in the general continuous time theory, can become quite technical; but, all of the main ideas and results carry through and in a similar form. The processes of concern will usually be either Markovian or perturbations of Markov processes, although more general processes will be dealt with in some of the examples. A vector-valued process $\{X_n\}$ is said to be *Markov* if, for any set A in its range space,

$$P\{X_{n+1} \in A \mid X_i, i \le n\} = P\{X_{n+1} \in A \mid X_n\}.$$

I.e., given the most recent state, the rest of the past does not provide any more information on the future. The process is said to have a time invariant transition function, if the right hand side does not depend on n. Markov processes are the natural stochastic analog of deterministic, dynamical processes (such as difference or differential equations) and are the most popular stochastic models in applications all through the physical, biological and social sciences.

Owing to the nature of the stability problem, it can be viewed as a robustness property of a system. If the basic model has a certain stability property; then, it is crucial in applications that a wide variety of perturbations of that model have the same or a very similar property since the precise nature of the disturbances, whether random or not, will rarely be known. The primary methods can deal with such perturbations. Non-Markovian correlated noise processes are important, particularly as "driving forces" in adaptive systems, and will be discussed as well.

2. Stability and Liapunov Functions

A quick review of the classical, Liapunov function method will be given first; since the various steps and interpretations motivate the methods for the stochastic case. Consider the ordinary differential equation $\dot{x} = f(x), x \in \mathbb{R}^r$, Euclidean *r*-space, and let $f(\cdot)$ be continuous. We wish to know whether $x(t) \to 0$ as $t \to \infty$ for any initial condition $x(0)$, and whether the solution stays close to zero for all time if it starts near zero. Furthermore, will these properties continue to hold under small perturbations of $f(\cdot)$? Keep in mind that although the smallness of a perturbation might guarantee that the solutions would remain close over any finite time interval; it is not guaranteed over the infinite time interval. These questions will be examined from the classical Liapunov function point of view.

Throughout the article, the function denoted by $V(x)$ will have the following properties; and, these properties *will not be repeated:* $V(\cdot)$ is a continuous, real-valued, nonnegative function on $\mathbb{R}^r$ which goes to infinity as $x \to \infty$. Also, $V(0) = 0$ and $V(x) > 0$ for $x \neq 0$. Other properties will be assumed from time to time. As used below, such $V(\cdot)$ will also be called *Liapunov functions*. The basic idea is that $V(\cdot)$ plays the role of a generalized energy function. If, with this measure of energy, the energy decreases or does not increase at each point of any path in an appropriate set, then conclusions of a stability nature can be drawn. The situation is analogous in the stochastic case. For positive λ, define the set $Q_\lambda = \{x : V(x) < \lambda\}$.

In the present differential equations case, suppose also that $V(\cdot)$ is continuously differentiable. The method is as follows. Compute the derivative of $V(\cdot)$ along a trajectory of the system, namely,

$$\frac{dV(x(t))}{dt} = V_x'(x(t)) f(x(t)) = -k(x(t)), \tag{1}$$

where the ' denotes *transpose* and which defines the continuous function $k(\cdot)$. For some $\lambda > 0$, let $x(0) \in Q_\lambda$ and suppose that $k(x) \geq 0$ for $x \in Q_\lambda$. Then, the following (typical Liapunov function) analysis holds. Since $dV(x(t))/dt \leq 0$ as long as $x(t) \in Q_\lambda$, the entire path stays in Q_λ for that initial condition. Furthermore, the function $V(x(t))$ is non-increasing; hence, it converges to some value $\bar{V}$. Since

$$V(x(t)) - V(x(0)) = -\int_0^t k(x(s)ds, \tag{2}$$

for all t, it follows that

$$0 \le \int_0^\infty k(x(s))ds < \infty \tag{3}$$

for all initial conditions $x(0) \in Q_\lambda$. The function $V(x(t))$ is said to have a *contraction property* in Q_λ. Since $x(t) \in Q_\lambda$ for all t, it follows that $x(\cdot)$ is uniformly continuous in t which, taken together with Eq. (3), implies that $x(t)$ converges to the set $K_\lambda = \{x : x \in Q_\lambda, k(x) = 0\}$. If $k(x) = 0$ implies that $x = 0$, then $x(t)$ converges to zero and we have the classical Liapunov function stability theorem. Furthermore, it can be shown that if the path starts near the origin; then, it will not stray far from it before eventually going to zero.

In many cases, $k(x) = 0$ at points x other than zero. Since, asymptotically, the trajectory hovers around K_λ and satisfies the differential equation in any case; it can be seen that $x(t)$ can only converge to a subset of K_λ which contains the full trajectory of the differential equation, starting from any point in it. This result is codified in the well-known *LaSalle invariance theorem*, which has an important stochastic analog, and will now be described.

A set G is called an *invariant set* for the ODE: if for each $x \in G$, there is a solution to the ODE on the doubly infinite time interval $(-\infty, \infty)$, which lies entirely in G and satisfies $x(0) = x$. The *invariant set theorem* states that if $x(0) \in Q_\lambda$ in which $k(x) \ge 0$, *then* $x(t)$ *converges to the largest invariant set in* K_λ. This theorem provides a substantial extension of the basic Liapunov stability theorem, and its analog for the stochastic problem is of considerable importance there as well.

Consider the following deterministic two dimensional example:

$$\begin{aligned}\dot{x}_1 &= x_2\\ \dot{x}_2 &= -x_1 - x_2,\end{aligned}$$

and set $V(x) = x_1^2 + x_2^2$, the energy. Then $\dot{V}(x) = -2x_2^2 = -k(x)$. Thus, $k(x) = 0$ tells us nothing about the component x_1, and the classical Liapunov function method cannot be used directly to show that $x(t) \to 0$. However, since $k(x) = 0$ implies that $x_2 = 0$ and $V(x(t))$ is still non-increasing, to find the limit invariant set, we need to find the set of points where $x_2 = 0$ and which supports a trajectory on $(-\infty, \infty)$. Unless $x(t) = 0$ for all t, the trajectory will eventually have nonzero $x_1(t)$. Thus, the origin is the only invariant set on which $k(x) = 0$. Consequently, $x(t) \to 0$. Clearly, the invariant set theorem simplifies and extends the basic Liapunov function method. The continuity condition on $k(\cdot)$ can be weakened, but care must be exercised.

A Discrete Time Liapunov Function Method

Suppose that the dynamical system is given in discrete time as $X_{n+1} = f(X_n)$, for a measurable function $f(\cdot)$. Suppose that, for some $\lambda > 0$,

$$V(X_{n+1}) - V(X_n) = -k(X_n) \leq 0 \qquad (4)$$

for $X_n \in Q_\lambda$. Then, if $X_0 \in Q_\lambda$, $X_n \in Q_\lambda$ for all n and X_n converges to the set $K_\lambda = \{x : k(x) = 0, x \in Q_\lambda\}$. There is a discrete time analog of the invariant set theorem, so that X_n actually converges to the largest invariant set in K_λ.

A major problem in the application of the method is the difficulty of finding suitable Liapunov functions, apart from some special but important classes. This is also true for the stochastic problem. In some cases, as the above example shows, obvious energy functions can be used. In other cases, as in adaptive control theory, there are natural “error functions”, which provide useful Liapunov functions. However, even in simple cases, where a useful Liapunov function is available, it can be used to study the stability under perturbations and time variations. Liapunov functions for deterministic problems can be used to study the effects of stochastic perturbations. For such problems, an important question is whether a stability property of the deterministic system will carry over or be approximable under stochastic perturbations. Then, one does not always need the have the Liapunov function in hand. Merely knowledge of its existence is enough, and the functions will exist if the system is appropriately stable.

The local “contraction” properties Eqs. (1) or (4) were crucial for the convergence. The key to the power of the Liapunov function approach is that purely local properties such as the sign of the derivative $V_x'(x)f(x)$ at each x can be used to get global properties of the trajectories. Such local contraction properties would occur for only trivial stochastic problems, where no matter what the realization of the random variables, there is a contraction at each step. However, there are some remarkable results in probability theory that allow us to get results that are close to the deterministic ones.

-
-
-

Bibliography

Benveniste A., Métivier M., and Priouret P. (1990). *Adaptive Algorithms and Stochastic Approximations*, 365 pp. Berlin and New York: Springer-Verlag. [An introduction to stochastic recursive algorithms.]

Breiman L. (1992). Probability Theory. *Classics in Applied Mathematics*, 7, 421 pp. Reprint of the 1968 Addison-Wesley edition. Philadelphia: SIAM. [A good book for the basics concerning stochastic convergence and martingales.]

Dai J.G. (1995). On Positive Harris Recurrence of Multiclass Queueing Networks: A Unified Approach Via Fluid Models. *Ann. Appl. Probab*., **5,** 49-77. [Typical of recent applications of stability to queueing networks.]

Loparo K.A. and Feng K.X. (1995). Stability of Stochastic Systems. *The Control Handbook*. (ed. W.S. Levine), pp., 1105-1126. Boca Raton: CRC Press. [A good discussion concerning Liapunov exponents, which are useful to get rates of convergence or divergence.]

Khasminskii R.Z. (1982). *Stochastic Stability of Differential Equations*. Amsterdam: Sijtho, Noordho, Alphen aan den Rijn. [A classical book on stochastic stability. Contains instability theorems also.]

Kurzweil J. (1963). On The Inversion of Liapunov's Second Theorem On Stability of Motion. *Amer. Math. Soc. Translations, Ser 2*, **24,** 19-77. [Gives broad conditions under which there exists a deterministic Liapunov function with given smoothness.]

Kushner H.J. (1967). *Stochastic Stability and Control*, 161 pp. New York: Academic Press. [A classical book on stochastic stability.]

Kushner H.J. (1971). *Introduction to Stochastic Control*, 390 pp. New York: Holt, Rinehart and Winston. [Contains a development of the discrete time theory of stochastic stability and the invariant set theorem.]

Kushner H.J. (1972). Stochastic Stability. *Stability of Stochastic Dynamical Systems: Lecture Notes in Math*, (ed R. Curtain), pp. 97-124. Berlin and New York: Springer-Verlag. [A comprehensive development of the continuous time stochastic stability theory, with a general prof. of the invariant set theorem.]

Kushner H.J. (1984). *Approximation and Weak Convergence Methods for Random Processes with Applications to Stochastic Systems Theory*, 269 pp. Cambridge, Mass: MIT Press. [Develops the perturbed Liapunov function methods.]

Kushner H.J. and Yin G. (1997). *Stochastic Approximation Algorithms and Applications*, 417 pp. Berlin and New York: Springer-Verlag. [A comprehensive book on all aspects of stochastic recursive algorithms, and the role of stability theory. Includes examples of the used of perturbed Liapunov functions.]

LaSalle J.P. and Lefschetz S. (1961). *Stability by Liapunov's Direct Method with Applications*, 134 pp. New York: Academic Press. [A classical book on deterministic stability.]

Ljung L. and Söderström T. S. (1983). *Theory and Practice of Recursive Identification*, 529 pp. Cambridge, Mass: MIT Press. [A good reference for recursive identification algorithms.]

Biographical Sketch

In the last 45 years, **Harold Kushner** has worked mainly on stochastic control. His nine books and over two hundred papers contain the seminal works for a substantial part of the field. These include stochastic stability (Markov and non-Markov), nonlinear filtering, distributed and delay systems, stochastic variational methods, stochastic approximation, efficient numerical methods for Markov chain models, the numerical methods of choice for general continuous time systems, singular stochastic control, stochastic networks, heavy traffic analysis of queueing/communications systems, wide band noise driven systems, problems with small noise effects, approximation methods, nearly optimal control and filtering for non-Markovian systems, and algorithms for function minimization. He received the IEEE field award in control, is a past chairman of the Applied Mathematics Dept. and the Lefschetz Center for Dynamical Systems (Brown University).