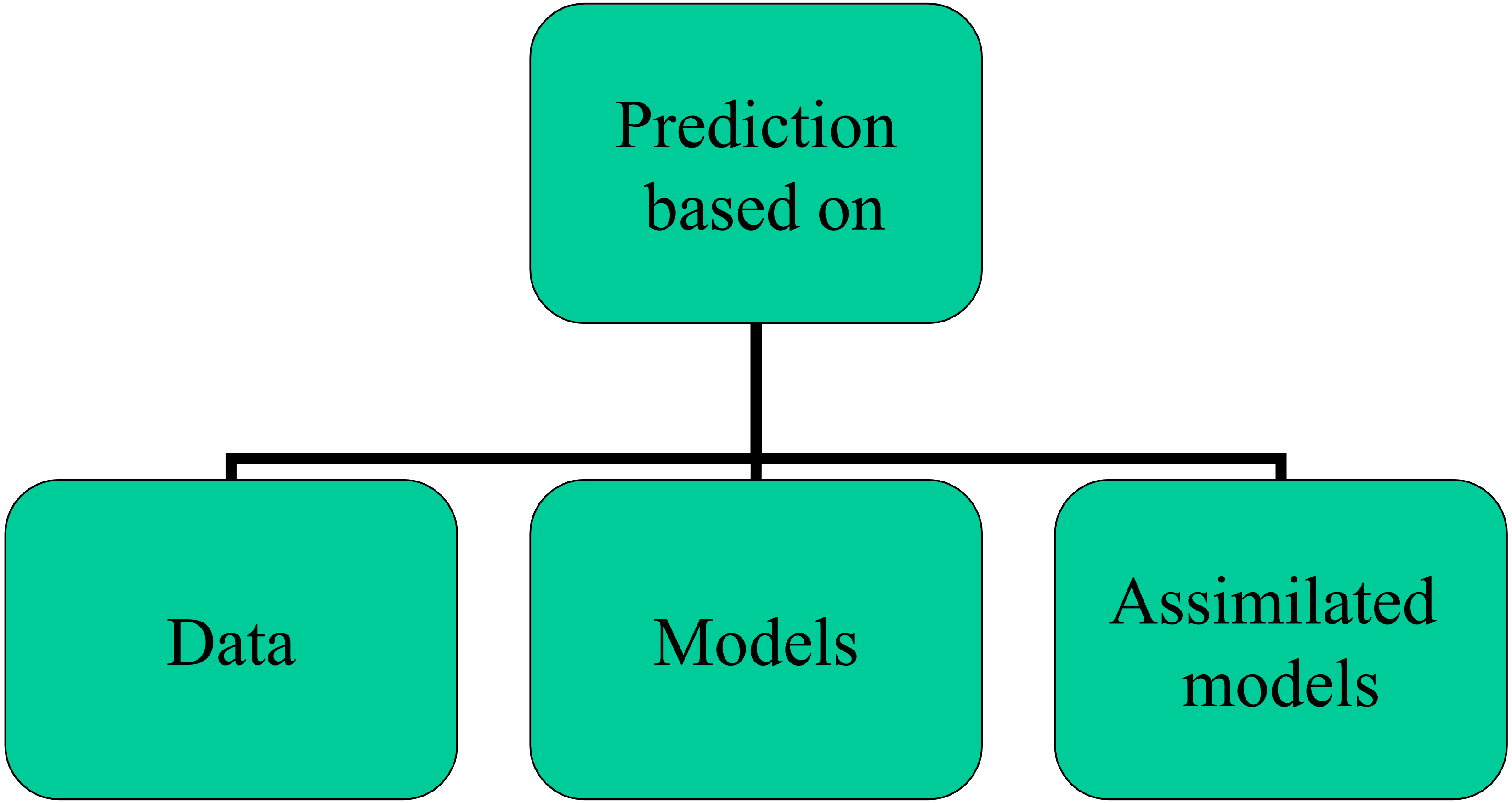
Predictability: an overview

S. Lakshmivarahan
School of Computer Science
University of Oklahoma
Norman, OK

Need for prediction is all pervasive

- Each of us one time or other have looked into our **crystal ball** to check/verify how tomorrow looks
- Many of what we do in life is driven by prediction of some sort
- **Electric Power** companies want to plan on generation depending on the predicted high temperature for tomorrow
- **Federal Reserve Board** adjusts the short-term interest rates based on the prediction of inflationary tendencies

Basis for prediction



Prediction based on historical data: Examples

- **Election results** based on polling data
- Price of an **old house/used car** based on "appraisal"
- **Stock prices** based on recent observations
- Prediction of "Tsunami" by "sea gypsies" based on the observed recession of the waters edge – There was a nice piece by Mr. Bob Simon in CBS 60 minutes program in mid August 05
- **Astrology** and Fortune telling
- This list goes on with many more examples

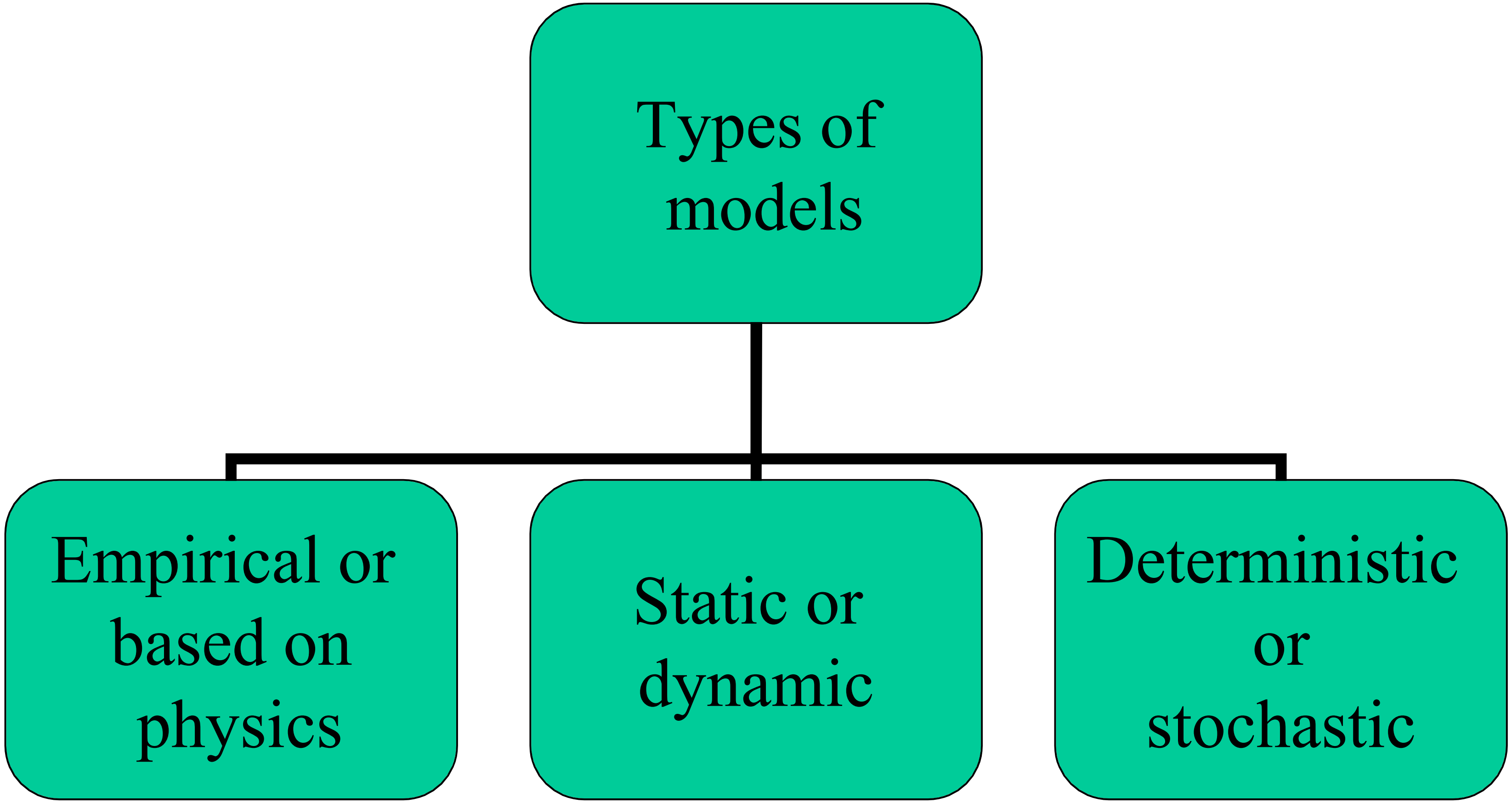
Why data based prediction?

- Data/Observations represent the true state of the system being observed
- Permits inductive generalization
- This is how progress in Science is made
- Early astronomers used this line of reasoning to create a wealth of knowledge
- Earth rotates around the sun in elliptic orbits
- Discovery of the law of gravitation
- Discovery of various models for the atom

Need for modeling data

- **Trade-off**: need large amount of data to have better accuracy in prediction but it gets to be more difficult to understand the inherent patterns in the data especially when the data set is large
- Motivates to create a model that can **generate/replicate** the “essence” observations modulo noise
- Example: Kepler derived the three laws that bears his name which was the basis for Newton to create the laws of gravitation

Nature and types of models: a classification



Empirical models: a classification

- **Time series modeling** –ARIMA models in Econometrics, Signal processing
- **Regression models** in Statistics
- **Data Mining**: principal component analysis
- **Supervised learning** using neural networks
- **Clustering** (unsupervised) to classify data into two various sub-classes

Physics based models

- **Inverse square law** of gravitation
- A class of **baratropic models** that describes the motion of a hurricane: PDE that describes the vortex - a cork in a moving fluid
- Motion of the **planets** in the solar systems (ODE)

Examples

- Inverse square law: deterministic/static
- Barotropic model: deterministic/dynamic
- Statistical regression: stochastic/static
- Time series: stochastic/dynamic

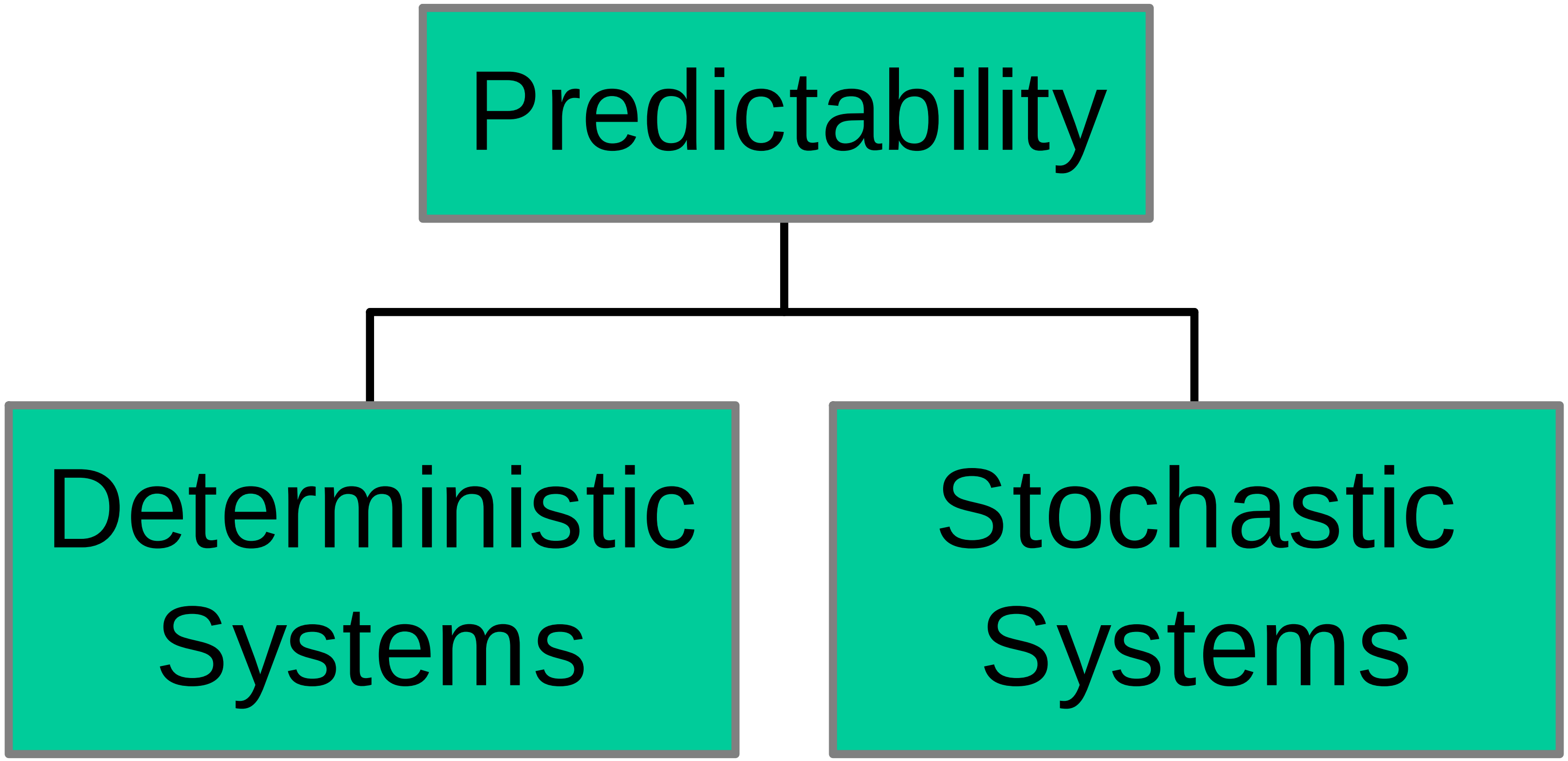
Instantiation of models using data assimilation

- Static models have **unknown parameters**
- Dynamic models have **unknown initial/boundary conditions/parameters**
- **Estimate** the unknowns of the model using well known **data assimilation** schemes
- Generate predictions using the **assimilated models**

From prediction to predictability

- Some events are **perfectly predictable** - solar eclipse
- Some events are predictable only for a **limited horizon** – winter weather is better than summer weather
- An astronomer can predict where the moons of Jupiter will be at 11.00pm this evening but has no idea where his teenage son/daughter will be.

Model based predictability analysis



Predictability of deterministic Systems

- From the time of Newton (1643-1727) to 1890 it was believed that deterministic dynamical systems are **perfectly predictable** in the sense that the states of a system can be computed for infinitely long time
- This belief rests on the fact if a system can be solved in closed form, then its solution is known for all times.

Dynamical systems

- **Linear systems**
- Behavior of linear system is completely catalogued
- Stable vs.unstable
- Stable: periodic or asymptotically stable
- Linear systems are perfectly predictable for all times

- **Nonlinear systems**
- Only a very small fraction of nonlinear systems can be solved in closed form and hence perfectly predictable
- For a vast majority of nonlinear systems numerical solution is the only recourse
- The question of predictability arise only with this latter class of systems

Limited predictability: The first encounter

- Poincaré between 1893-1912 for the first time encountered the possibility of **limited predictability** while trying to numerically solve a simple 3-body problem for which closed form solution was (and is still) not known

3-body problem

- Poincaré for the first time found out that very **small changes in the initial conditions resulted in large changes in trajectories** during the numerical integration
- This result sowed the seeds for the development of the theory of deterministic chaos as we know it today

Qualitative Theory of differential equations

- Deals with the question: can we predict the long term qualitative behavior of the solution of dynamical systems without actually solving the equations?
- **Lyapunov theory** of stability paved the foundations for this line of attack
- It provides sufficient conditions for stability by analyzing the behavior of certain functional – called a **Lyapunov function**, along the trajectory of the system

Qualitative Theory

- The period 1900-1960 witnessed great advances in the development of the qualitative theory that laid the foundations of the stability theory as we know it today
- V.V. Nemytskii and V.V. Stepanov (1960) **Qualitative Theory of Differential Equations**, Princeton University Press, 523 pages
- E. Lorenz (1963) "Deterministic non-periodic flows", **Journal of Atmospheric Sciences**, Vol 20, 130-141.
- This paper by Lorenz extensively uses the concepts and results from the above book

Early 1950's

- The question of meteorological predictability came to the forefront around 1950's - thanks to the new emerging digital computer era, when large scale numerical weather prediction became feasible
- P.D. Thompson(1957)"Uncertainty of initial state as a factor in the predictability of large scale atmospheric flow patterns", **Tellus,** Vol 9, 275-295.
- This paper by Thompson is the first in recent times to analyze the limited predictability due to uncertainties in the initial field using a stochastic framework

Foundations of limited predictability

- Ed Lorenz in 1963 revived the interest in this area by the (accidental) invention of the new phenomenon of **deterministic chaos** while numerically integrating a system of nonlinear equations –known known as the Lorenz's model
- Chaos is defined as the **sensitivity to** small changes in **initial conditions** as suspected by Poincare some seven decades ago
- Systems that exhibit sensitivity to initial conditions by definition have only **limited predictability**

Special aspects of nonlinear systems

- Like linear systems nonlinear systems can be asymptotically stable, periodic and unstable
- Unlike linear systems, the specialty however is that a nonlinear system can be unstable and yet be bounded- known as chaos
- Chaos is like an angry energetic tiger in a cage – bounded and yet not predictable

Estimation of rate of growth of small errors: basic idea

Let

$$\dot{x} = f(x)$$

be the given model

Let

$$X_{k+1} = M(X_k)$$

be the discrete time version of the model

Then

$$e_{k+1} = D_M(x_k)\, e_k$$

gives the **dynamics of first order perturbation** (TLM) where $D_M(x_k)$ is the **Jacobian** of M at x_k

Let for $t \geq s \geq 0$. Then the **state transition matrix** is given by

$$D_M(t:s) = D_M(x_t)D_M(x_{t-1})\ldots\ldots..D_M(x_s)$$

The error at time (k+1) is related that at time 0 is then given by

$$e_{t+1} = D_M(t:0)\ e_0$$

The **Rayleigh coefficient** has been used to analyze the rate of growth of errors:

$$r_{t+1} = (e_{t+1}{}^T e_{t+1}) / (e_0{}^T e_0)$$

$$= (e_0{}^T D_M{}^T(t:s) D_M(t:s) e_0) / (e_0{}^T e_0)$$

The **eigen values** of $D_M{}^T(t:s)D_M(t:s)$ which are called the **singular values** of $D_M\,(t:s)$ determine the rate of growth in finite time (t-s)

Asymptotic rate of growth of small errors

Renormalization strategy:

$$\lambda = \lim_{N\to\infty} \lim_{e_0\to 0} (1/N) \Sigma \log (e_{k+1}{}^T e_{k+1}/e_k{}^T e_k)$$

This number is called the leading **Lyapunov exponent** where the summation is from 0 to N

It denotes the average of the **logarithm of the amplification** along the trajectory

A good reference for the computation of Lyapunov exponents is:

T. S. Parker and L.O. Chua (1989) **Practical Numerical Algorithms for Chaotic Systems**, Springer Verlag

Typical values of Lyapunov exponent

Steady state	Attractor set	Lyapunov exponent
Equilibrium point	Point	$\lambda < 0$
Periodic orbit	cycle	$\lambda = 0$
Chaotic	Fractal structure	$\lambda > 0$

If λ is positive, then small errors grow at the rate give by

$$e^{\lambda t}$$

The **predictability limit** is

$$t_p = 1/\lambda$$

Which is the time required for the initial errors to grow by a factor e which is analogous to the notion of time constant

Note: Only for simplicity in presentation we have assumed that there is only one positive exponent. In practice there could be more. In such a case we take the sum of all the positive exponents. In such a case the predictability limit would be even lesser than when there is only one positive exponent

Note: The number of Lyapunov exponents is equal to the dimension of the system

For classic Lorenz system of 1963 the three exponents are given by $\lambda_1 = 0.9$, $\lambda_2 = 0$, and $\lambda_3 = -12.8$

In a recent paper

E. N. Lorenz (2005) " A look at some of the details of he growth of initial uncertainties", **Tellus**, Vol 57A, 1-11

provides a very good discussion of predictability of yet another related 3 dimensional system by computing the three Lyapunov exponents

Summary

- Limited time (deterministic) predictability is an exclusive property of nonlinear deterministic systems for which there is no known closed form solution and whose properties have to be understood only using extensive computer simulations
- The values of the Lyapunov exponent is a clear indicator of predictability limit

Stochastic differential equations (SDE) : Ito equation

Let

$$dX_t = f(X_t, \alpha)dt + \sigma(x_t)dw_t$$

be the given dynamics with x_0 is the initial condition, the forcing dw_t is the standard Brownian motion or Wiener process and α is the parameter and $\sigma(X)$ is the state dependent covariance matrix

State vector is x =$(x_1, x_{2,}, \ldots x_n)$, f=$(f_1, f_2, \ldots f_n)$,

w=$(w_1, w_2, \ldots w_n)$

Markov diffusion process

- Solution of the SDE is a Markov process
- This provides a natural extension of ODE to SDE
- Fokker-Planck (1900's) Kolmogorov (1930's) - derived equations for the evolution of the probability density of Markov process
- Ito calculus provides the frame work for the analysis of functionals of this class of Markov process

Source of randomness in models

Randomness can enter in three ways:

- Initial/boundary condition – otherwise deterministic system
- Random forcing
- Random coefficients
- Combinations of these three sources

Predictability in stochastic systems

- Given the pdf of the sources of randomness – initial condition or the random forcing or of parameters, drive the dynamical equations for the evolution of the PDF $P_t(X_t)$ of the state X(t)
- Given any subset B of the model space, then we can compute the probability that X_t belongs to the given subset B
- The subset B could denote the normal course of events or extreme events, etc
- Predictability of a stochastic system is completely solved in principle

Random initial conditions

- Let $P_0(X_0)$ be the PDF of the initial state X_o of the model: $\sigma(x) = 0$
- Once a random realization of IC is picked, the system acts as though it is a purely a deterministic system
- The PDF $P_t(X_t)$ of X_t is the given by the classical Liouville's equation:

Randomness only in the initial condition: Liouville's equation

- Let $P_t(X_t)$ be the probability density of the state x_t at time t. The P_t is given by the solution of the Liouville's equation which is a (parabolic) partial differential equation where $P_o(x_o)$ is the given initial distribution

$$\partial P_t / \partial t + \sum_i \partial [f_i P_t] / \partial x_i = 0$$

- For a derivation of Liouville’s equation refer to the following sources:
- T. L. Saaty (1967) **Modern nonlinear equations**, McGraw Hill, Chapter 8
- T.T. Snoog (1973) **Random differential equations in Science and Engineering**, Academic Press

Random IC and random forcing: Kolmogorov forward or Fokker-Planck equation

P_0 is the initial condition

$$\partial P_t / \partial t = -\sum_i \partial[f_i P_t] / \partial x_i$$

$$+ (1/2) \sum_{i,j} \partial^2 \{[\sigma(x)\sigma^T(x)]_{i,j} P_t\} / \partial x_i \partial x_j$$

- For a derivation of Kolmogorov's forward and backward equations refer to the following sources:
- A. H. Jazwinski (1970) **Stochastic Process and Filtering**, Academic Press

Relation to Liouville's equation

- Kolmogorov's forward equation reduces to Liouville's equation if there is no random forcing term: when the coefficient $\sigma(x)$ of dw_t is zero
- Notice the natural nesting of the results
- For methods for solving Kolmogorov's equations refer to:

  H. Risken (1984) **The Fokker-Planck Equation**: **methods of solutions and applications**, Springer Verlag series in Synergetics, Vol 18, 454 pages

Effect of data assimilation on predictability in stochastic models

- The evolution of the probability density of the state of the stochastic dynamic system with data assimilation called Stratonovich-Kushner-Zakai (S-K-Z) equation was first derived in the non-linear filtering literature
- Normalized version of this equation was first derived by S-K in the early 1960's and the unnormalized version by Z in the late 1960's

S-K-Z equation

$$dX_t = f(X_t, \alpha)dt + \sigma(x_t)dw_t \qquad \text{: Model}$$

$$dz_t = h(x_t)dt + dv_t \qquad \text{: Observation}$$

dw_t and dv_t two uncorrelated Wiener processes

$$dP_t = -\sum_i \{\partial[f_i P_t]/\partial t\}dt$$

$$+ \tfrac{1}{2}\sum_{i,j}\partial^2 \{[\sigma(x)\sigma^T(x)]_{ij}P_t\}/\partial x_i \partial x_j \, dt$$

$$+ P_t \, [dz_t - E(h(x_t)dt]^T \, \Sigma^{-1} \, [h(x_t) - E(h(x_t)]$$

Σ is the covariance of observation noise

Nesting of equations:

- S-K-Z is a stochastic PDF (random initial conditions, random forcing and random observations)
- When there is no observation, S-K-Z reduces to Kolmogorov's forward equation
- When there is no random forcing, it further reduces to Liouville's equation

Reference to S-K-Z equations

H.J. Kushner (1962) “On the differential equations satisfied by the conditional densities of Markov processes with applications”, **SIAM Journal on Control and Optimization**, Vol 2, 106-119

- M. Zakai (1969) “On the optimal filtering of diffusion processes”, **Zeitschrift fur Wahrschienlickeitstheorie und Verwanddte Gebiete**, Vol 11, 230-243
- G. Kallianpur (1980) **Stochastic Filtering Theory**, Academic Press

N. Wiener (1949) **Extrapolation, Interpolation and Smoothing of Stationary Time Series with Engineering Applications**, Wiley (originally published as a classified document in Feb,1942)

A.N. Kolmogorov (1941) **Interpolation, extrapolation of Stationary Random sequences,** Bulletin of Academy of Sciences, USSR, Series on Mathematics, Vol 5 (Translation by RAND corporation Memorandum RM -3090 April, 1962)

R. E. Kalman (1960) “ A new approach to linear filtering and prediction”, **Transactions of the American Society of Mechanical Engineeting**, **Journal of Basic Engineering**, Series –D, Vol 82, 35-45

R. S. Bucy and P.D. Joseph (1968) **Filtering for stochastic processes with applications to guidance**, Interscience publications

R.S. Bucy (1994) **Lectures on discrete filtering**, Springer Verlag

- Solution of S-K-Z equation:

- J. F. Bennaton (1985) “Discrete time Galerkin approximations to nonlinear filtering”, **Journal of Mathematical Analysis and Applications**, Vol 110, 364-383

Approximate moment dynamics

- Since S-K-Z type equations are not easy to solve, great attention has been given to quantifying the evolution of the first **K-moments** of the random state x(t)
- This line of argument was pursued by Thompson (1957), non-linear filtering literature
- **Extended Kalman filters** (1960), Epstein (1960's) – **stochastic dynamics**
- A major challenge facing this approach is the classical **moment closure** problem

Reference for moment closure

- A. H. Jazwinski (1970) **Stochastic Process and Filtering Theory**, Academic Press
- E. S. Epstein (1969) **Stochastic Dynamic Prediction**, Tellus, Vol XXI, 739-759

Summary

- Thanks to the developments in nonlinear filtering theory, predictability of stochastic dynamic system has been solved in principle. The challenge is to solve the S-K-Z equations

Conclusion

- Determination of predictability limit of large scale (such as those that arise in meteorology) deterministic and stochastic dynamical systems provides some of the most demanding computational problems of our time

References

J.Lewis, S. Lakshmivarahan and S. k. Dhall (2006) **Dynamic Data Assimilation: a least squares approach**, Cambridge University Press (Chapters 31 & 32 are on predictability)

Interest in predictability is not limited to meteorology

Clive W. J. Granger and O. Morgenstern (1970) **Predictability of Stock Market Prices**, MIT Press

A. Cutler (1983) **Justice and predictability**, Macmillan (Punishment in crime deterrence)

H. H. Stevenson (1998) **Do lunch or be lunch; the power of Predictability in creating your own future**, Harvard Business School Press, Boston, Mass (Business forecasting)