Prediction, bad prediction, optimal prediction

Alexandre Chorin
University of California, Berkeley

Argument: Predictions (in economics, climate, physics) are often very poor because:

- bad models,
- insufficient data,
- unreasonable mathematical assumptions.

Example 1: Space travel.

A rocket is sent to towards the moon, one wants to predict its position at later times.

Model: Newton's (or Einstein's) equations of motion.

Data: Initial position, thrust, geophysical and astronomical data.

The equations of motion: force= mass X acceleration.

acceleration= rate of change of velocity, velocity= rate of change of position,

summary: rate of change of state= function of the state.

Example 2: Weather forecasting

model: general circulation model; uncertainties in tropical dynamics, cloud physics, mountains, hills, forests, turbulence, chaos, numerical approximation.

data: sparse, partial, uncertain observations.

other examples: climate prediction, economic forecasting.

Linear vs. nonlinear

In a linear model, the effect is proportional to the cause.

Nothing breaks. Structures cannot collapse. Hurricanes do not happen, the stock market never crashes.

We are interested in nonlinear models.

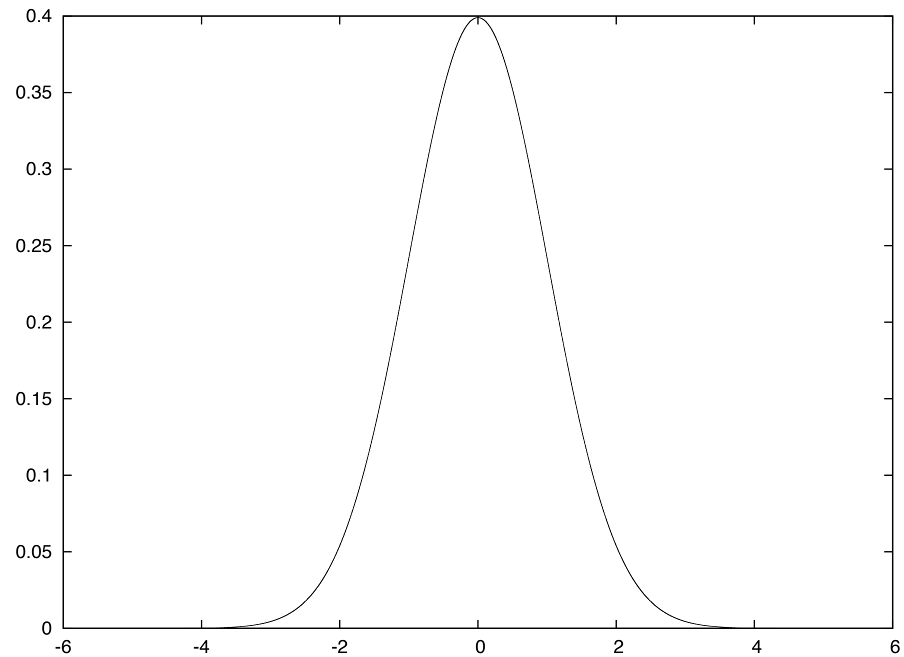
Random variable A quantity whose value is determined by experiment. You may not know this value before the experiment, but you know in advance the probabilities of the various outcomes.

Expected value (mean, average). The sum of the various outcomes multiplied by their probabilities. If a random variable takes on the value 5 with probability 1/2 and the value 3 with probability 1/2, it expected value is 5/2+3/2=4.

Before the experiment, the expected value is the best estimate of the outcome.

Given partial information, the best estimate of the outcome is the conditional expectation (conditional mean, conditional average), which is the average over those outcomes that are compatible with the partial data.

Example of a random variable: Gaussian



<u>”Central limit theorem”:</u> the sum of many independent random variables is approximately Gaussian.

<u>”Law of large numbers”:</u> the average of many outcomes approximates the expected value.

<u>Sums</u>. Random variables add up more slowly than non-random quantities: The sum of n Gaussian variables is of the order of $\sqrt{n}$.

<u>Computer sampling</u>. One can sample random variables on the computer: one can generate numbers so that the probability of getting any specific value matches a prescribed probability.

How does one make a forecast in the presence of uncertainty?

step 1: make a model with some of the variables random, in particular, make some assumptions about the uncertainties. The probabilities of the random variables are obtained from past experience.

step 2: sample the random variables that appear; now the model is no longer random. Make a prediction.

step 3: do step 2 many times, and average the results. What you get is the best estimate of the future given what you know or do not know.

Sampling is NOT the same as replacing a variable by its best estimate. Example: you earn an random interest with mean 10% and variance .1 for 30 years. The best estimate of what you will have after 30 years with an investment of 1 dollar is $ 89.98; if you replace the interest by its mean you get only $ 20.08.

First idea: the models have the form:

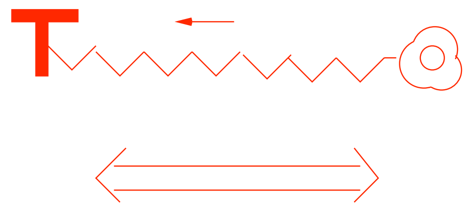
Rate of change of a state= a function of the state.

Pick a subset of variables to solve for, and approximate the function of the state by its conditional expectation given the values of these variables. (i.e., replace the right hand side of the equation by the best approximation available to you).

Used in turbulence, plasma physics,...

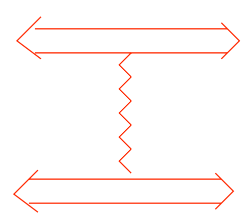
Example:

A harmonic oscillator is a system that consists of a pebble, a rubber band, and a nail:

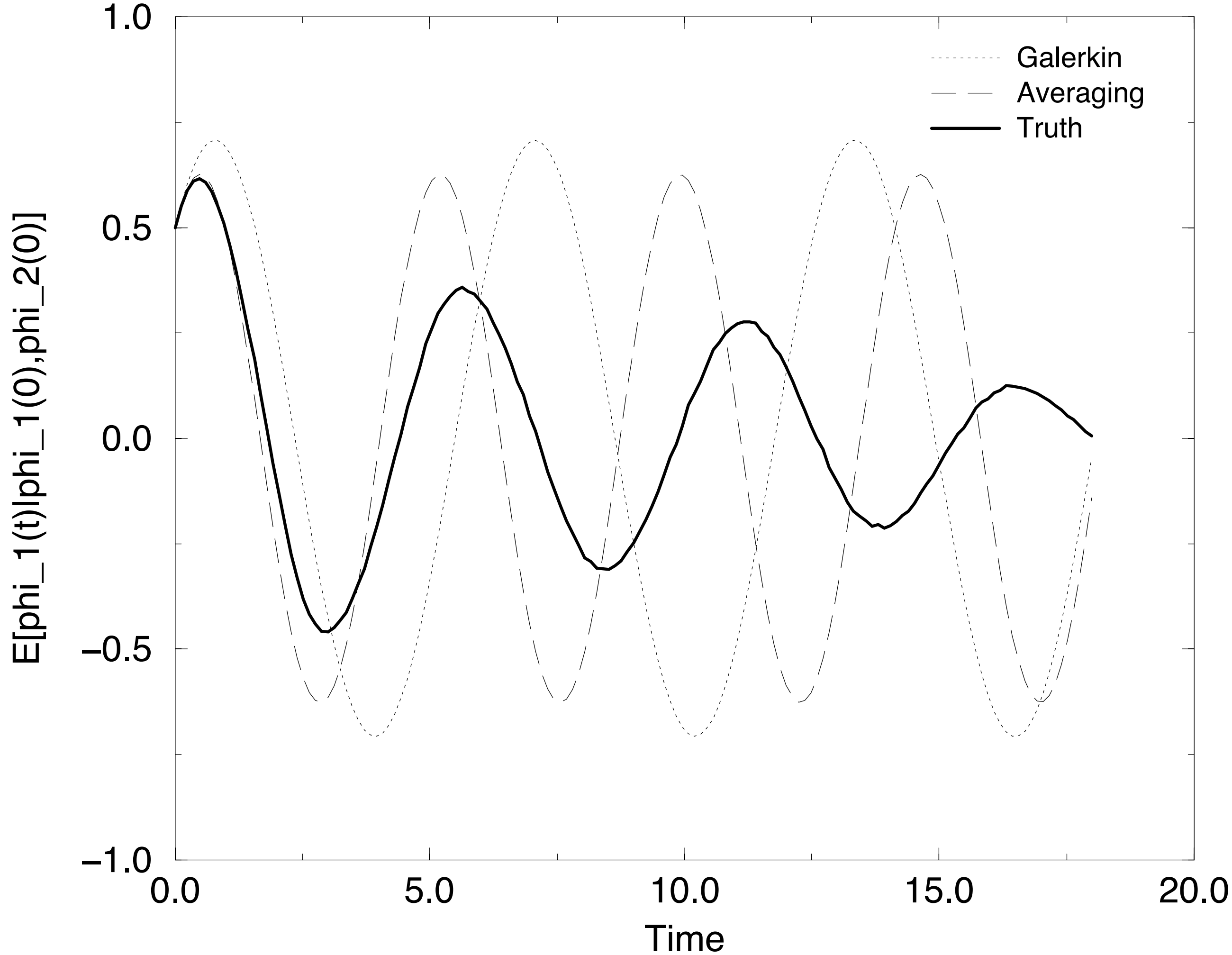


Harmonic oscillator

Consider the "Hald model": two harmonic oscillators, coupled by a nonlinear rubber band, with initial data for one, and a probability density for the initial values of the other.



"Hald model"


Galerkin
Averaging
Truth
E[phi_1(t)|phi_1(0),phi_2(0)]
1.0
0.5
0.0
−0.5
−1.0
0.0
5.0
10.0
15.0
20.0
Time

Fallacies: Average of squares $\neq$ square of averages:

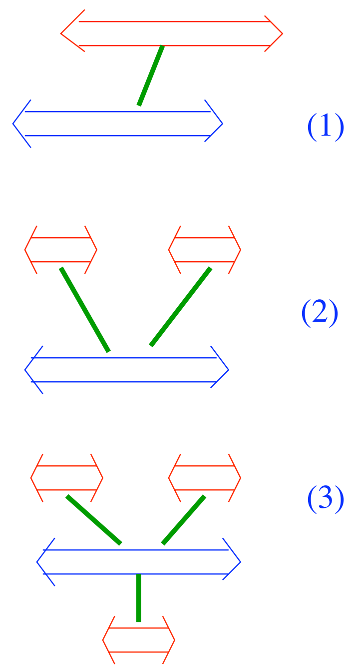
1	5	square of average: 9
1	25	average square: 13.

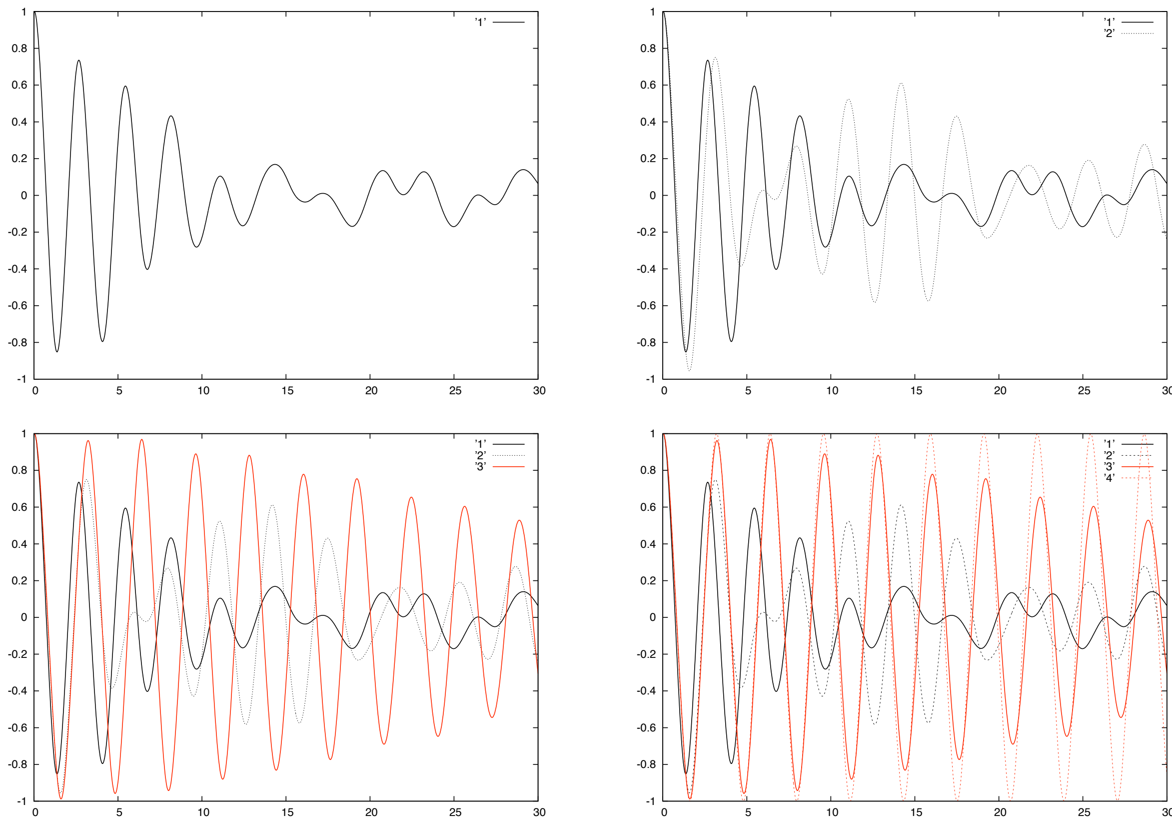
Similarly: The solution of the average equation is not the average solution (solving equations and averaging do not commute).

True equation: rate of change of best estimate= average of right-hand- side + memory term !

memory term= average interaction with fluctuations.

Approximation of memory term: assume everything you do not know is small and fast ("separation of scales").




'1'
'1'
'2'
'1'
'2'
'3'
'1'
'2'
'3'
'4'