* 1. Stochastic Process
* Def. 2.36. A stochastic process is a family of random variables, , indexed by a real parameter and defined on a common probability space .

%%% Kim’s comment

A stochastic process (or random process) is a time varying random variable, i.e., for any fixed , the process is a random variable.

%%%

* Ex. 2.37

Consider a sinusoidal signal whose amplitude and frequency is a R.V. as

Hence an experiment with a realization

Let another experiment with a realization

And so on. Now for a fixed time ,

The collection of all experiments

which is a R.V. with mean

* Def. 2.38.

1. A stochastic process is said to be **continuous** in probability at t if

for all

1. **Skip**: A stochastic process is said to be separable if there exists a countable, dense set such that for any closed set

differ by a set such that

1. Skip: Theorem 2.40. **The rational numbers** in provide a separating set S.

* Def. 2.42. Let X be a random process defined on the time interval, T. Let be a partition of the time interval, T. If the increments, are mutually independent for any partition of T, then X is said to be a process with **independent increments**.

%%% Kim’s comment

By the definition, the increment , which are R.V.s are independent, but may not be independent. %%%

* Def. 2.43 We say that a random process, X, is a Gaussian process if for every finite collection, the corresponding density function,

is a Gaussian density function.

%% Kim’s comment

Hence the pdf is an form

Where

a constant so that

the mean of

covariance of i.e.,

%%%

* Def. 2.44 We say that a random process X is a Gaussian process if every finite linear combination of the form

is a Gaussian random variable

* Def 2.45. A random process, where T is a subset of the real line, is said to be a **Markov process** if for any increasing collection

or, equivalently

%%% Kim’s comment

Consider a deterministic signal

Then

which means for any , , which means it is sufficient that only one value of is needed.

In Markov, the probability conditioned on the previous history data is the same to the probability conditioned on only one previous data.

%%%

* 1. Gauss-Markov Processes – **The fundamental**

1. Dynamics

* State , is a known matrix, is a Gaussian Random sequence.
* Given Conditions

1. Noise

where

1. The initial states : the Gaussian with mean , the variance

And the initial is independent of the noise,

1. The mean and covariance

The mean satisfies the dynamics as

1. The covariance satisfies the dynamics
   1. Non-linear Stochastic Difference Equations 🡪 skip (later…)

**3. Conditional Expectations and Discrete-Time Kalman Filtering**

* 1. Minimum Variance Estimation
* Problem statement – static parameter estimation

measurement noise

Find or estimator to minimize the error in some sense

* Biased / un-Biased

If , the estimator is an un-biased estimator.

* The minimum variance estimator
* Remark 3.1.argmin

1. Without measurement , the minimum variance estimator is the mean of
2. With the measurement Minimum variance (conditional) estimator:
3. The minimum variance estimator is the conditional mean or the mean

* Difference : linear least square estimator

Consider measured 3 data points as

If data is modeled as a linear line

Here, is unknown to be estimated. Define the error as

To find the best estimator is to minimize the sum of errors as

Sol: The error sum is written as a matrix form

i.e.,

the best least square error estimator is

**In Machine Learning it is called linear regression problem.**

**The difference between least square and min mean variance estimator is the knowledge of the noise characteristics, i.e. regression has no information on the noise however, min mean variance should have**.

%%%%

* Theorem 3.6. Given the equation (3.1). **if the estimate is a function of , then the minimum variance estimate is the conditional mean. i.e.,**
* **One property : is un-biased**

Since by the iterative expectation

* Convolution: pdf of z which is sum of two independent RVs.

Consider two random variables which are independent with pdf as Define

Find pdf of Z

Solution:

Now

( or respect to y, )

* This integral is called the convolution integral.
* Example

The sum of two X,Y, . Find

1. Convoiultion

Since is plotted in the following

z-1

z

x

-1

y

Hence with the different , is

z-1

z

x

z-1

z

x

z-1

z

x

z-1

z

x

1. ,

1. In summary
2. Not using convolution from the joint pdf.

Since two are independent,

1. We may see in (x,y) co-ordinate

X

Y

z = x+y

* Ex.3.8: conditional mean and variance of the sum of two RV’s

Let measurement , , are independent. Find the the minimum variance estimate

* a prior information,
* a noise,
* In order to find , so need to find

Since , need

1. First find

First the noise pdf is

Hence

%% HA\_Week\_4.1

1. Plot
2. Find

%%%

1. Find

By Bayesian law, . Hence should be evaluated.

Since ,

%% warning : it is followed not by convolution but by linear sum. %%

Now each interval corresponding to

* ,

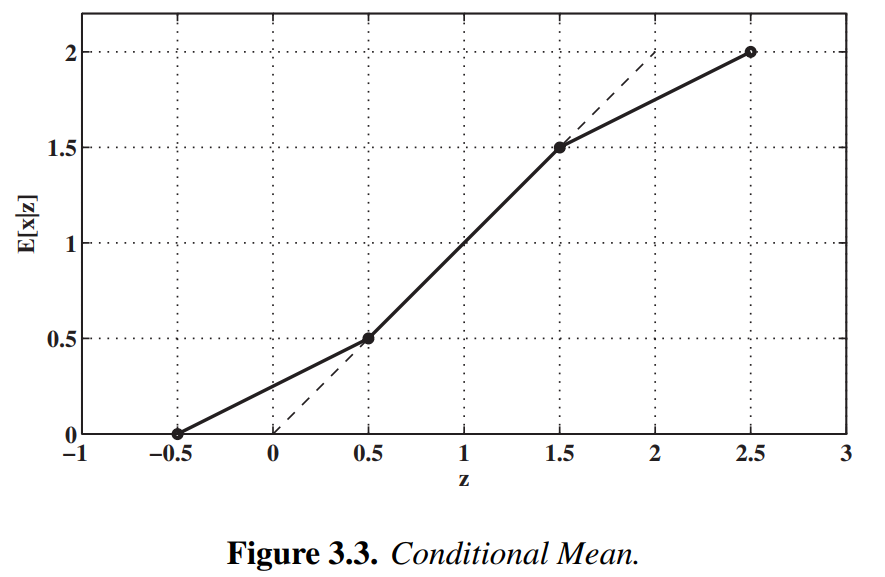
Hence

* since ,
* , since ,

1. Compute the min variance estimator,

Since

2. not defined.



%% Analysis: the dashed line is the linear estimator, . The best estimator in the sense of MV is 7/4 if .

* 1. Conditional Estimate of a Gaussian Random Vector with Additive Gaussian Noise
* Modelling
* %% I will notify means is independent.
* Problem: find

To get the solution, we need

1. **Since are Gaussians, is a Gaussian 🡪 need to be prove**
3. **, and it is a Gaussian. 🡪 need to be proved**
4. How to find given ?

* Solution

where

* How to find ?

First find then . The pdf is already known in theorem 2.30 as

* Find : **He use a trick**. Define a new RV as

%% Tip.

%%%

* By assumption

And

And

And

so that

And

Rearrange as

Here

From this equation, identify

And the covariance

%% Remark:

* The **MV estimator** in (3.14) is .

Also due to the “iterated expectation rule”, it is unbiased, i.e.,

* The covariance does **not depend on the measurement .** it depends on the variances of input and the measurement matrix
* (3.14) is called **“Kalman filter”.** The weighting on the difference between the measurement and the average of the to be estimated is called **“Kalman Gain”**
  + 1. Processing Measurements Sequentially
* Batch and recursive process

Example: a random process as

Find the average of

1. Batch Process
2. Recursive process

Define

Then

1. Merits for recursive way

* The memory size is lower than the batch type
* The result can be acquired at every step, which is more informative

1. The static Kalman (3.14) is a batch process. In dynamic system, it is more efficient to calculate the Kalman gain at the sample time.

* Recursive method for Kalman gain

In (3.2),

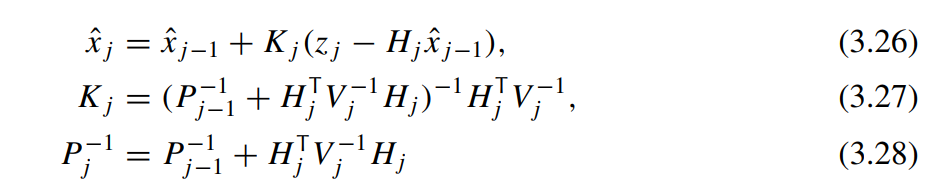
The measurement will be sequentially sampled

1. The covariance of in Batch
2. The covariance in Recursive

Define

Then

Hence by induction



* Procedure:
  1. Assumption:

1. The measurement matrix and variance of the noise are constant
2. The variance of state and noise are
   1. Initialization: Before the initial measurement, initialize is needed

Initial condition on the state:

Initial condition of estimator covariance:

1. Recursive computing

For , before measurement

Measurement . Substitute

For , before measurement

Measurement .

And so on,..