REF. NO. KF # 5 (Parts 1 and 2)

1

STOCHASTIC ESTIMATION

The Bayesien Approach to Parameter Estimation

Part 1

2

- MOTIVATION

In most physical or socioeconomic problems we make actual experiments to increase our knowledge about "something" of interest.

3

- The very act of making a measurement in the real world with a real "sensor" is fundamentally an uncertain process (recall the Heissenberg uncertainty principle in quantum mechanics). Hence, measurements are inherently unreliable or noisy.

4

- Sometimes we can directly measure the quantity or variable of interest (subject to an inherent measurement error associated with the sensor).

5

- However, there are cases that the measurements that we can, or care to, carry out contain only indirectly information about the quantity or variable of interest (still subject to measurement errors).

6

PRIOR VS POSTERIOR INFORMATION

- From a philosophical point of view it is reasonable to suppose that before we make an experiment we know something about the variable(s) of interest. If we knew nothing at all why should we be interested in the variables?

  We shall refer to such knowledge before the experiment as prior information (no matter how bad it may be).

7

- It is also reasonable to suppose that we know the physical significance of our measurement with respect to the variable of interest.

  Example: The physical measurement of the dimensions of Ms. Smith contains no information about the height of Mr. Brown, but it may contain information about his blood pressure!

8

- Thus, before we make the experiment we have another type of prior information namely the relation between the measurements that we may carry out and the quantities that we are interested in.

9

- Another type of prior information pertains to the accuracy of our "sensor." The sensor accuracy may be known from "manufacturing data", or from previous carefully controlled experiments (independent of the one that we are about to carry out) on the sensor(s) themselves.

  Remark: Often an increase in sensor accuracy requires more monetary expenditures.

10

THE VALUE OF DOING EXPERIMENTS

- Intuitively, if we expend time, money, and energy to carry out an experiment, we should certainly hope that after the experiment we should know more about the quantity of interest than before. Thus,

$$\boxed{\text{Posterior information} \geq \text{Prior Information}}$$

- Irrelevant experiments should preserve the equality about prior and posterior information.

11

Remark: If we had absolutely no information about the quantity of interest, then we would be hard pressed to

(a) establish the relevance of the experiment.

(b) the role of the sensor accuracy.

12

MATHEMATICAL MODELLING

- Variables or quantities of interest

$$x_1, x_2, \ldots, x_n \quad \leftarrow \text{real scalars}$$

Notation:

$$\underline{x} \triangleq \begin{bmatrix} x_1 \\ x_2 \\ \cdots \\ x_n \end{bmatrix}$$

RANDOM parameter vector of interest

13

- Actual variables that we measure

$z_1, z_2, \ldots, z_m$ real scalars

Notation:

$$\underline{z} = \begin{bmatrix} z_1 \\ z_2 \\ \cdots \\ z_m \end{bmatrix}$$ measurement vector

Each z_i contains the measurement error.

14

- Actual errors introduced by measurements

$\theta_1, \theta_2, \ldots, \theta_r$ usually, $r = m$

$$\underline{\theta} = \begin{bmatrix} \theta_1 \\ \theta_2 \\ \cdots \\ \theta_r \end{bmatrix}$$ measurement noise vector

15

- Relation of experiment to quantities of interest

$$\underline{z} = \underline{g}\,(\underline{x}, \underline{\theta}) \qquad (1)$$

i.e. $z_k = g_k(x_1, \ldots, x_n, \theta_1, \ldots, \theta_r)$ $k = 1, 2, \ldots, m$

Remark: Knowledge of the mapping (function) $\underline{g}(\cdot,\cdot)$ forms part of our prior information.

16

- MODELLING OF UNCERTAINTY ON VECTOR $\underline{x}$ BEFORE THE EXPERIMENT.
- $\underline{x}$ is modelled as a random vector.
- Prior information on $\underline{x}$ is modelled by the assumption that its probability density function, $p(\underline{x})$, is known

$p(\underline{x})$: Prior PDF of $\underline{x}$

$$p(\underline{x}) = p(x_1, x_2, \ldots, x_n)$$

scalar valued function of many variables

17

<u>MODELLING OF SENSOR ERRORS (i.e. MEASUREMENT UNCERTAINTY) PRIOR TO EXPERIMENT</u>

- Measurement noise vector $\underline{\theta}$ is modelled as a random vector.
- Manufacturers specification on sensor accuracy (part of prior information) are modelled by assuming that the probability density function, $p(\underline{\theta})$, is available

$$p(\underline{\theta}) = p(\theta_1, \theta_2, \ldots, \theta_r)$$

$P(\underline{\theta})$: prior PDF of $\underline{\theta}$, the sensor noise vector

18

- <u>Common Assumption</u>: The inherent uncertainty of sensors has nothing to do with the inherent uncertainty on our prior knowledge of the parameter vector $\underline{x}$.
- Mathematically this is modelled by assuming that $\underline{x}$ and $\underline{\theta}$ are <u>independent</u>

$$\rightarrow p(\underline{x}, \underline{\theta}) = p(\underline{x}) \cdot p(\underline{\theta})$$

19

<u>THE UNCERTAINTY ON THE MEASUREMENT VECTOR $\underline{z}$</u>

- From eq. (1) ,

$$\underline{z} = \underline{g}(\underline{x}, \underline{\theta})$$

- We can see that <u>before</u> the experiment we do not know what the measurements will be. Since $\underline{x}$ and $\underline{\theta}$ are random vectors, $\underline{z}$ (prior to the experiment) will be also a random vector, with prior probability density function

$$p(\underline{z}) = p(z_1, z_2, \ldots, z_m)$$

$P(\underline{z})$: prior PDF of measurement vector

20
REMARKS

- Under our assumptions the probability density function $p(\underline{z})$ can be evaluated
- After the experiment, our "sensors" have measured $\underline{z}$; hence it is no longer random.
  $\underline{z}$

21
MODELLING OF POSTERIOR INFORMATION ON $\underline{x}$

In general, since

- $\underline{x}$ was uncertain to begin with
- the sensor measurements were uncertain

we would expect that after the measurement we would not still know the value of $\underline{x}$ perfectly (but only "better"). Hence after the measurement the parameter vector of interest is still a random vector.

22

- Parameter vector of interest prior to experiment:

$$\underline{x} \rightarrow p(\underline{x})$$

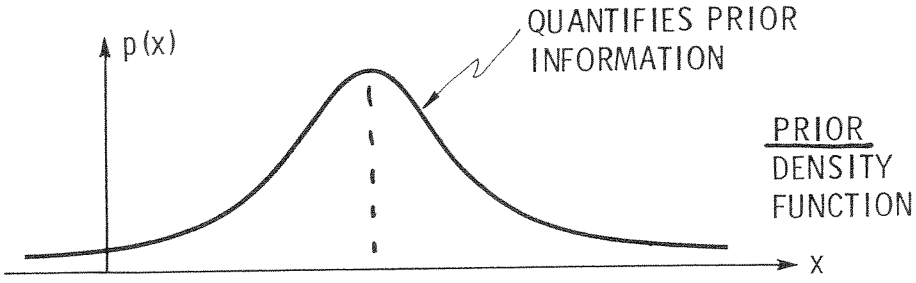
- Parameter vector of interest after the experiment:

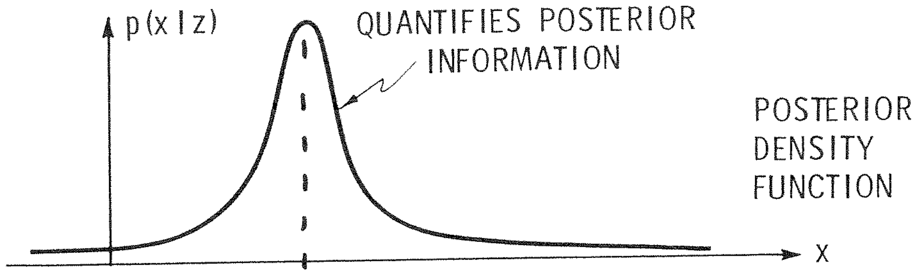
$$\underline{x}/\underline{z} \rightarrow p(\underline{x}/\underline{z})$$

Read: $\underline{x}|\underline{z} \Rightarrow$ "$\underline{x}$ given $\underline{z}$"

The conditional density function (or posterior density function) $p(\underline{x}/\underline{z})$ models the inherent uncertainty on $\underline{x}$ persisting through the measurement.

23





- Different mean
- Smaller variance

24

- <u>COMPUTATION OF POSTERIOR DENSITY FUNCTION</u>
- Main tool is <u>Bayes Rule</u>

$$p(\underline{x}, \underline{z}) = p(\underline{x}/\underline{z})\, p(\underline{z}) \tag{2}$$

$$p(\underline{x}, \underline{z}) = p(\underline{z}/\underline{x})\, p(\underline{x}) \tag{3}$$

25

- Interested in

$$\boxed{p(\underline{x}/\underline{z}) = \frac{p(\underline{x}, \underline{z})}{p(\underline{z})} = \frac{p(\underline{z}/\underline{x})\, p(\underline{x})}{p(\underline{z})}} \tag{4}$$

<u>Computational question</u>

- $p(\underline{x})$ was assumed known

Need to evaluate

- $p(\underline{z}/\underline{x})$ and $p(\underline{z})$

26

EVALUATION OF $p(\underline{z}/\underline{x})$ → If I knew $\underline{x}$, what can I say about $\underline{z}$ before I measure it?

In fundamental measurement relation

$\underline{z} = \underline{g}(\underline{x}, \underline{\theta})$; $\underline{z} \in R_m$, $\underline{x} \in R_n$, $\underline{\theta} \in R_r$

$\underline{x}$ is now viewed as given (no longer a random parameter

Hence, assumption that sensor measurements are always in error, because of measurement noise ⇒

$\underline{z}$ is random vector whenever
$\underline{\theta}$ is random vector

27

- But

m = number of sensed variables = number of sensors

r = number of measurement noises

To have an one-to-one correspondence between number of sensors and number of measurement noises, we must have

$$\boxed{r = m} \tag{5}$$

standing assumption from now on

28

- It is also reasonable to assume that any two different values of measurement noise do not yield the same measurement

$\underline{z} = \underline{g}(\underline{x}, \underline{\theta}) \Rightarrow \underline{\theta} = \underline{g}^{-1}(\underline{x}, \underline{z})$

i.e.

$\underline{g}(\underline{x}, \; .)$ is one-to-one and onto (6)

29

Under the two assumptions (5) and (6)
$p(\underline{z}/\underline{x})$ can be evaluated by

$$p(\underline{z}/\underline{x}) = \frac{1}{\det J_\theta} p(\underline{\theta}) \qquad (7)$$

where J_θ is the mxm (r=m) Jacobian matrix *

$$J_\theta = \frac{\partial \underline{g}(\underline{x}, \underline{\theta})}{\partial \underline{\theta}}, \underline{x} \text{ is parameter} \qquad (8)$$

Hence $p(\underline{z}/\underline{x})$ can be evaluated from the prior knowledge of $p(\underline{\theta})$ and $\underline{g}(\underline{x}, \underline{\theta})$

30

EVALUATION OF $p(\underline{z})$

- In principle, $p(\underline{z})$ can be evaluated analytically using the relations

$$\underline{z} = \underline{g}(\underline{x}, \underline{\theta})$$

$$p(\underline{x}, \underline{\theta}) = p(\underline{x})\, p(\underline{\theta})$$

- General formulas are complex we shall see how this is done in special cases

* $\frac{\partial \underline{g}(\underline{x}, \underline{\theta})}{\partial \underline{\theta}}$ is the mxm Jacobian matrix of partial derivatives with respect to $\underline{\theta}$

$$ij\text{-th element} = \left[\frac{\partial \underline{g}(\underline{x}, \underline{\theta})}{\partial \underline{\theta}}\right]_{ij} = \frac{\partial g_i(x_1, \dots, x_n, \theta_1, \dots, \theta_m)}{\partial \theta_j}$$

$i, j = 1, 2, \dots, m$

Example

$$z_1 = g_1(x_1, x_2, \theta_1, \theta_2) = x_1^3 \theta_1 + x_2 \theta_1 \theta_2 + x_2^2 \theta_2^2$$

$$z_2 = g_2(x_1, x_2, \theta_1, \theta_2) = (\sin x_2)\theta_1 + \theta_2$$

$$\frac{\partial \underline{g}(\underline{x}, \underline{\theta})}{\partial \underline{\theta}} = \left[\begin{array}{c|c} x_1^3 + x_2\theta_2 & x_2\theta_1 + 2x_2^2\theta_2 \\ \hline \sin x_2 & 1 \end{array}\right]$$

31

STOCHASTIC ESTIMATION

The Bayesian Approach to Parameter Estimation

Part 2

32

USEFUL STATISTICS

- The prior and posterior probability density functions provide the most general mathematical description for the uncertainty.

33

- For engineering problems one is often interested in a more "summary" type of information pertaining to the parameter vector of interest. Two intuitively appealing "statistics" are
  (a) A "good" estimate of $\underline{x}$, denoted by $\hat{\underline{x}}$.
  (b) A "good" measure of the estimation error

$$\tilde{\underline{x}} \triangleq \underline{x} - \hat{\underline{x}}$$

$\wedge$ denotes estmate

$\sim$ denotes estimation error

34

- COMMON ESTIMATES OF RANDOM VECTORS

Let $\underline{y}$ be a random vector, with probability density function $p(\underline{y})$. Then, the most often used estimates are:

1) The mean

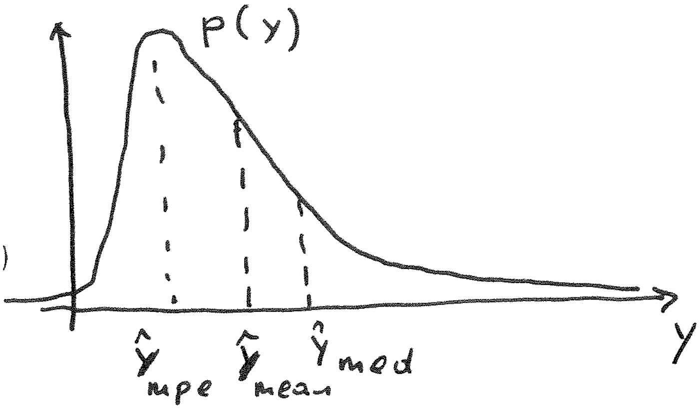
$$\hat{\underline{y}}_{mean} = E\{\underline{y}\} = \int_{-\infty}^{+\infty} \cdots \int \underline{y}\, p(\underline{y})\, d\underline{y} \qquad (9)$$

35

2) <u>The median</u>, i. e. the estimate that minimizes the maximum possible magnitude of the estimation error. If we let $\hat{\underline{y}}_{med}$ denote this estimate, then

$$\| \underline{y} - \hat{\underline{y}}_{med} \| \le \max \| \underline{y} - \hat{\underline{y}} \| \qquad (10)$$

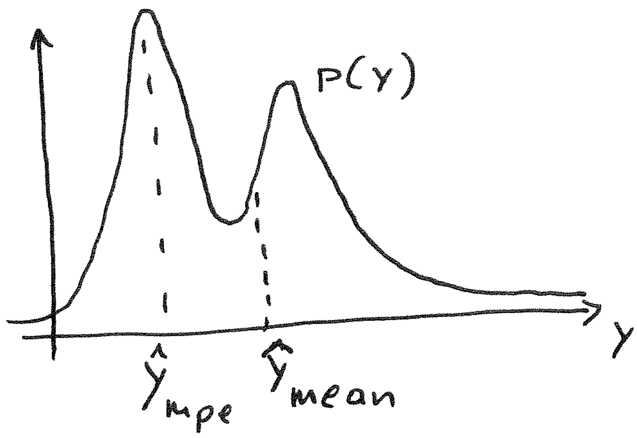
where $\hat{\underline{y}}$ is any other estimate.



36

3) <u>The most probable estimate</u> (maximum likelihood in Bayesian sense), $\hat{\underline{y}}_{m.p.e.}$ which corresponds to the highest peak of the density function, i. e.

$$p\left(\hat{\underline{y}}_{m.p.e.}\right) \ge p(\underline{y}) \quad \text{for all } \underline{y} \qquad (11)$$



37

Given any estimate $\hat{\underline{y}}$, we can get an idea of the confidence of that estimate by computing the matrix

$$\underline{\Sigma} = E\left\{\left(\underline{y}-\hat{\underline{y}}\right)\left(\underline{y}-\hat{\underline{y}}\right)'\right\} \qquad (12)$$

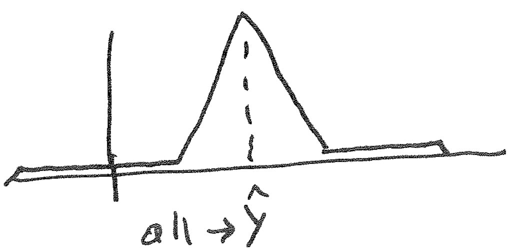
• If $\hat{\underline{y}} = E\{\underline{y}\}$, then $\underline{\Sigma}$ is the covariance matrix

38

• <u>FACT</u> Any single mode symmetric probability density function, has the property that

$$\hat{\underline{y}}_{mean} = \hat{\underline{y}}_{median} = \hat{\underline{y}}_{m.p.e.} \qquad (13)$$

← may not be gaussian, e.g.



39

• FACT The Gaussian probability density function with mean $\underline{\bar{y}}$ and covariance matrix $\underline{\Sigma}$, where $\underline{y} \in R_n$

$$\boxed{p(\underline{y}) = (2\pi)^{-\frac{n}{2}} (\det \underline{\Sigma})^{-\frac{1}{2}} \cdot \exp\left\{-\frac{1}{2} (\underline{y}-\underline{\bar{y}})' \underline{\Sigma}^{-1} (\underline{y}-\underline{\bar{y}})\right\}} \qquad (14)$$

Shorthand:
$\underline{y} \sim N(\underline{\bar{y}}, \underline{\Sigma})$

is a single mode symmetric probability density function, so that

$$\hat{\underline{y}}_{mean} = \hat{\underline{y}}_{median} = \hat{\underline{y}}_{m.p.e.} = \underline{\bar{y}} \qquad (15)$$

40

APPLICATION OF BAYESIAN ESTIMATION TO A LINEAR-GAUSSIAN EXAMPLE

• Definitions

$\underline{x} \in R_n$ parameter vector of interest

$\underline{z} \in R_m$ measurement vector

$\underline{\theta} \in R_m$ noise vector

• Measurement Equation (Linear)

$$\boxed{\underline{z} = \underline{H}\underline{x} + \underline{\theta}} \qquad (16)$$

$\underline{H}$ mxn known deterministic matrix

41

• PRIOR INFORMATION

• $\underline{x}$ is Gaussian random vector

$$\underline{x} \sim N(\underline{x}_b, \underline{\Sigma}_b)$$

$$\boxed{p(\underline{x}) = (2\pi)^{-\frac{n}{2}} (\det \underline{\Sigma}_b)^{-\frac{1}{2}} \cdot \exp\left\{-\frac{1}{2}(\underline{x}-\underline{x}_b)' \underline{\Sigma}_b^{-1} (\underline{x}-\underline{x}_b)\right\}} \qquad (17)$$

subscript "b" means "before experiment", i.e. prior

42

- $\underline{\theta}$ is Gaussian random vector

$$\underline{\theta} \sim N(\underline{0}, \underline{\Theta})$$

$$\boxed{p(\underline{\theta}) = (2\pi)^{-\frac{m}{2}} (\det \underline{\Theta})^{-\frac{1}{2}} \cdot \exp\left\{-\frac{1}{2}\underline{\theta}' \underline{\Theta}^{-1}\underline{\theta}\right\}} \qquad (18)$$

- $\underline{x}$ and $\underline{\theta}$ are independent

$\Longrightarrow p(\underline{x}, \underline{\theta}) = p(\underline{x})\, p(\underline{\theta})$

43

Prior (i. e. Before the Experiment) Statistics

$$\hat{\underline{x}} = E\left\{\underline{x}\right\} = \underline{x}_b \quad \text{(Prior Mean)} \qquad (19)$$

Prior Covariance

$$= \text{cov}\left[\underline{x};\underline{x}\right] = \underline{\Sigma}_b \qquad (20)$$

44

COMPUTATION OF POSTERIOR DENSITY FUNCTION $p(\underline{x}/\underline{z})$

- From Bayes rule

$$\boxed{p(\underline{x}/\underline{z}) = \frac{p(\underline{z}/\underline{x})\, p(\underline{x})}{p(\underline{z})}} \qquad (21)$$

- $p(\underline{x})$ is known - see eq. (17)

45

- Evaluation of $p(\underline{z}/\underline{x})$

$$\underline{z} = \underline{H}\,\underline{x} + \underline{\theta} \qquad (22)$$

Viewing $\underline{x}$ as known, then $\underline{H}\underline{x}$ is viewed as a deterministic vector. All uncertainty in $\underline{z}$ is caused by uncertainty in $\underline{\theta}$

recall: $\underline{z} = \underline{H}\,\underline{x} + \underline{\theta}$

46

Hence ($\underline{z}/\underline{x}$) is Gaussian   Think of $\underline{x}$ as given!

$$E\left\{\underline{z}/\underline{x}\right\} = \underline{H}\,\underline{x} + \underbrace{\cancel{E\left\{\underline{\theta}\right\}}}_{\underline{0}} = \underline{H}\,\underline{x} \qquad (23)$$

conditional mean of $\underline{z}/\underline{x}$

$$\mathrm{cov}\left[\underline{z};\underline{z}/\underline{x}\right] = \underline{\Theta} \qquad (24)$$

$$\boxed{p(\underline{z}/\underline{x}) = (2\pi)^{-\frac{m}{2}} (\det \underline{\Theta})^{-\frac{1}{2}} \cdot \exp\left\{-\frac{1}{2}(\underline{z}-\underline{H}\,\underline{x})' \,\underline{\Theta}^{-1} (\underline{z}-\underline{H}\,\underline{x})\right\}} \qquad (25)$$

47

EVALUATION OF $p(\underline{z})$ ← Prior!

- $\underline{z} = \underline{H}\,\underline{x} + \underline{\theta}$

→ $\underline{z}$ is Gaussian

$$E\left\{\underline{z}\right\} = \underline{H}\,E\left\{\underline{x}\right\} + \underbrace{\cancel{E\left\{\underline{\theta}\right\}}}_{\underline{0}} = \underline{H}\,\underline{x}_b \qquad (26)$$

48

$$* \quad \mathrm{cov}\left[\underline{z};\underline{z}\right] = \underline{H}\,\mathrm{cov}\left[\underline{x};\underline{x}\right]\underline{H}' + \mathrm{cov}\left[\underline{\theta};\underline{\theta}\right] = \underline{H}\,\underline{\Sigma}_b\,\underline{H}' + \underline{\Theta} \qquad (27)$$

because $\underline{x}$ and $\underline{\theta}$ are independent!

Hence

$$\boxed{p(\underline{z}) = (2\pi)^{-\frac{m}{2}} \cdot \left[\det\left(\underline{H}\,\underline{\Sigma}_b\,\underline{H}' + \underline{\Theta}\right)\right]^{-\frac{1}{2}} \cdot \exp\left\{-\frac{1}{2}\left(\underline{z}-\underline{H}\,\underline{x}_b\right)'\left(\underline{H}\,\underline{\Sigma}_b\,\underline{H}' + \underline{\Theta}\right)^{-1}\left(\underline{z}-\underline{H}\,\underline{x}_b\right)\right\}} \qquad (28)$$

Recall: if $\underline{w} = \underline{A}\,\underline{x} + \underline{B}\,\underline{y}$; $\underline{x}, \underline{y}$ independent random vectors

$$\mathrm{Cov}\left[\underline{w};\underline{w}\right] = \underline{A}\,\mathrm{cov}\left[\underline{x};\underline{x}\right]\underline{A}' + \underline{B}\,\mathrm{cov}\left[\underline{y};\underline{y}\right]\underline{B}'$$

49

THE POSTERIOR DENSITY p(x/z)

Subscript "a" means "after the experiment", i.e. posterior

- Substitution and linear algebra (lots!) yield structure of $p(\underline{x}/\underline{z})$ - i.e. plug in (30) to (32) into (29) and crank out.
- Structure of $p(\underline{x}/\underline{z})$

$\underline{x}|\underline{z} \sim N(\underline{x}_a, \underline{\Sigma}_a)$

$$\boxed{p(\underline{x}/\underline{z}) = (2\pi)^{-\frac{n}{2}}\left(\det\ \underline{\Sigma}_a\right)^{-\frac{1}{2}} \cdot \exp\left\{-\frac{1}{2}\left(\underline{x}-\underline{x}_a\right)'\underline{\Sigma}_a^{-1}\left(\underline{x}-\underline{x}_a\right)\right\}} \qquad (29)$$

$\underline{x}/\underline{z}$ is still Gaussian! ← this falls out of math! It is not an assumption

50

POSTERIOR MEAN $\underline{x}_a$

$$\boxed{\underline{x}_a = \underline{x}_b + \underline{\Sigma}_a\ \underline{H}'\ \underline{\Theta}^{-1}\left[\underline{z} - \underline{H}\underline{x}_b\right]} \qquad (30)$$

$\underline{x}_a = E\left\{\underline{x}/\underline{z}\right\}$ = optimal posterior estimate of $\underline{x}$ (31)

- $\underline{\Sigma}_a$ is posterior covariance matrix

51

POSTERIOR COVARIANCE $\underline{\Sigma}_a$

$$\boxed{\underline{\Sigma}_a = \underline{\Sigma}_b - \underline{\Sigma}_b\underline{H}'\left(\underline{H}\,\underline{\Sigma}_b\underline{H}' + \underline{\Theta}\right)^{-1}\underline{H}\,\underline{\Sigma}_b} \qquad (32)$$

$\underline{\Sigma}_a = \text{cov}\left[\underline{x};\underline{x}/\underline{z}\right]$ = posterior covariance matrix of $\underline{x}$ (33)

52

$$\boxed{\underline{\Sigma}_a^{-1} = \underline{\Sigma}_b^{-1} + \underline{H}'\ \underline{\Theta}^{-1}\ \underline{H}} \qquad (34)$$

$\Longrightarrow \underline{\Sigma}_a^{-1} \geq \underline{\Sigma}_b^{-1} \Longrightarrow \underline{\Sigma}_a \leq \underline{\Sigma}_b$ (35)

$\Longrightarrow$ reduced uncertainty

If $\underline{A}$ and $\underline{B}$ are positive definite matrices; $\underline{A} > \underline{0}$, $\underline{B} > \underline{0}$. Then

$\underline{A} \geq \underline{B} \Rightarrow \underline{x}'(\underline{A}-\underline{B})\underline{x} \geq 0, \forall \underline{x}$

$\underline{A} \leq \underline{B} \Rightarrow \underline{x}'(\underline{A}-\underline{B})\underline{x} \leq 0, \forall \underline{x}$

Numerical Example

$$p(x) \sim N(0,1) \Rightarrow p(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} \qquad (1)$$

$$p(\theta) \sim N(0,2) \Rightarrow p(\theta) = \frac{1}{\sqrt{2\pi}\cdot\sqrt{2}} e^{-\frac{\theta^2}{4}} \qquad (2)$$

$$z = x + \theta \qquad (3)$$

Suppose we measure $z = 1/2$ (4)

$$p(z|x) \sim N(x,2) \Rightarrow p(z|x) = \frac{1}{\sqrt{2\pi}\sqrt{2}} e^{-\frac{1}{4}(z-x)^2} \qquad (5)$$

$$E\{z\} = \cancel{E\{x\}}^{0} + \cancel{E\{\theta\}}^{0} = 0 \qquad (6)$$

$$var[z] = \overset{0}{var}[x] + \overset{0}{var}[\theta] = 1 + 2 = 3 \qquad (7)$$

$$\therefore\ p(z) \sim N(0,3) \Rightarrow p(z) = \frac{1}{\sqrt{2\pi}\sqrt{3}} e^{-\frac{z^2}{6}}$$

From (4)

$$p(x|z=\tfrac{1}{2}) = \frac{p(z|x)\,p(x)}{p(z)} = \frac{\frac{1}{\sqrt{2\pi}\sqrt{2}} e^{-\frac{1}{4}(\frac{1}{2}-x)^2} \cdot \frac{1}{\cancel{\sqrt{2\pi}}} e^{-\frac{x^2}{2}}}{\frac{1}{\cancel{\sqrt{2\pi}}\sqrt{3}} e^{-\frac{1}{24}}}$$

$$= \frac{1}{\sqrt{2\pi}\sqrt{\frac{2}{3}}} e^{\left[\underbrace{-\frac{1}{4}(\frac{1}{2}-x)^2 - \frac{1}{2}x^2 + \frac{1}{24}}_{a}\right]}$$

Can show that: $a = -\frac{3}{4}\left(x^2 - \frac{1}{3}x + \frac{1}{36}\right) = -\frac{3}{4}\left(x - \frac{1}{6}\right)^2$

$$= -\frac{1}{2}\,\frac{(x - 1/6)^2}{(2/3)}$$

$$\therefore\ p(x|z=\tfrac{1}{2}) \sim N\left(\frac{1}{6}, \frac{2}{3}\right) = N(x_a, \Sigma_a)$$

$$x_a = 1/6, \quad \Sigma_a = \frac{2}{3}$$

which can also be verified from eqs. (30) and (32)

53

NUMERICAL EXAMPLE

- Parameter vector $\underline{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$ (36)
- Prior Mean: $\underline{x}_b = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$ (37)
- Prior Covariance: $\underline{\Sigma}_b = \begin{bmatrix} 4 & 1 \\ 1 & 2 \end{bmatrix}$ (38)

54

- MEASUREMENT

$$z = x_1 + 3x_2 + \theta = \qquad (39)$$

$$\underbrace{\begin{bmatrix} 1 & 3 \end{bmatrix}}_{\underline{H}} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \theta$$

- $E\left\{\theta\right\} = 0$, $\operatorname{cov}\left[\theta ; \theta\right] = \Theta = 2$ (40)
- Numerical outcome: $z = 9$ (41)

55

- CALCULATION OF POSTERIOR COVARIANCE $\underline{\Sigma}_a$

$$\underline{\Sigma}_a = \underline{\Sigma}_b - \underline{\Sigma}_b \underline{H}' \left(\underline{H} \underline{\Sigma}_b \underline{H}' + \underline{\Theta}\right)^{-1} \underline{H}\underline{\Sigma}_b \qquad (42)$$

$$\underline{\Sigma}_b \underline{H}' = \begin{bmatrix} 4 & 1 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} 1 \\ 3 \end{bmatrix} = \begin{bmatrix} 7 \\ 7 \end{bmatrix} \qquad (43)$$

$$\underline{H} \underline{\Sigma}_b \underline{H}' + \underline{\Theta} = 28 + 2 = 30 \qquad (44)$$

56

$$\underline{\Sigma}_a = \begin{bmatrix} 4 & 1 \\ 1 & 2 \end{bmatrix} - \begin{bmatrix} 7 \\ 7 \end{bmatrix} \frac{1}{30} \begin{bmatrix} 7 & 7 \end{bmatrix}$$

$$= \begin{bmatrix} 4 & 1 \\ 1 & 2 \end{bmatrix} - \begin{bmatrix} 1.633 & 1.633 \\ 1.633 & 1.633 \end{bmatrix}$$

$$= \begin{bmatrix} 2.366 & -0.633 \\ -0.633 & 0.366 \end{bmatrix} \qquad (45)$$

57

CALCULATION OF POSTERIOR MEAN

$$\underline{x}_a = \underline{x}_b + \underline{\Sigma}_a \underline{H}' \underline{\Theta}^{-1} \left[\underline{z} - \underline{H} \underline{x}_b\right] \qquad (46)$$

$$= \begin{bmatrix} 1 \\ 2 \end{bmatrix} + \begin{bmatrix} 2.366 & -0.633 \\ 0.366 & 0.366 \end{bmatrix} \begin{bmatrix} 1 \\ 3 \end{bmatrix} \frac{1}{2}$$

$$\underbrace{\left[9 - \begin{bmatrix} 1 & 3 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \end{bmatrix}\right]}_{2}$$

$$= \begin{bmatrix} 1 \\ 2 \end{bmatrix} + \begin{bmatrix} 0.466 \\ 0.466 \end{bmatrix} = \begin{bmatrix} 1.466 \\ 2.466 \end{bmatrix}$$