

Chapter 5 Optimal Estimation

Part 1

5.1 Random Variables, Processes and Transformation



Outline

- 5.1 Random Variables, Processes and Transformation
 - 5.1.1 Characterizing Uncertainty
 - 5.1.2 Random Variables
 - 5.1.3 Transformation of Uncertainty
 - 5.1.4 Random Processes
 - Summary



Outline

- 5.1 Random Variables, Processes and Transformation
 - <u>5.1.1 Characterizing Uncertainty</u>
 - 5.1.2 Random Variables
 - 5.1.3 Transformation of Uncertainty
 - 5.1.4 Random Processes
 - Summary



5.1.1 Characterizing Uncertainty

- Uncertainty =
 - Not Known: Bias, scale systematic error
 - E.g. Temperature sensitivity
 - Not Knowable: Noise, randomness, unpredictability
 - E.g. "drift"
- Fact of life:
 - Some randomness is fundamental
 - It can't be measured.
- Humans do a good job coping...



Modeling Uncertainty

- An oxymoron?
- Distributions are models.
- Algebraic and differential equations are models.
- We can "pass distributions through" equations to get other distributions.
 - one point at a time, or...
 - as a complete distribution



Carnegie Mellon THE ROBOTICS INSTITUTE

5.1.1.1 Types of Uncertainty

• We usually consider it to be additive:

$$x_{meas} = x_{true} + \varepsilon \quad \hat{x} = x_{true} + \varepsilon \quad \text{estimate}$$

- ϵ may be zero, a constant, or a function of anything.
- ε may be:
 - Systematic (="deterministic")
 - Random (= "stochastic")
 - a combination of both.

Random error is called **unbiased** of it has a mean of zero.

 Most of all ε is unknown. Otherwise we would take it out.

5.1.1.1 Real and Ideal Signals

• Below: bias, scale errors, and two "outliers".



Carnegie Mellon THE ROBOTICS INSTITUTE

5.1.1.1 Real and Ideal Signals

• Below: Saturation, nonlinearity, deadband





5.1.1.1 Real and Ideal Signals

Might model the errors like so:

 $\varepsilon = Q + \partial \theta + N(\mathbf{0}, \mathbf{O})$

- Note the appearance of model parameters of both kinds:
 - systematic (a,b)
 - stochastic (μ , σ)



THE ROBOTICS IN

Removing errors

- Systematic \rightarrow calibration:
 - Fit a line to the last graph
- Stochastic \rightarrow filtering
 - Smooth out the wiggles
- Correlation \rightarrow differential measurement
 - Reject the common component

Know Your Model

• You can fit a line to anything.



Carnegie Mellon THE ROBOTICS INSTITUTE

Know Your Model

• You can fit a line to anything.



Carnegie Mellon THE ROBOTICS INSTITUTE

Outline

- 5.1 Random Variables, Processes and Transformation
 - 5.1.1 Characterizing Uncertainty
 - <u>5.1.2 Random Variables</u>
 - 5.1.3 Transformation of Uncertainty
 - 5.1.4 Random Processes
 - Summary



Probability as Frequency Distribution

- Events that occur randomly may nonetheless have a knowable probability distribution.
- Its not unusual to know the distribution but never be able to perfectly predict an individual event.
- Knowing one distribution allows you to compute others.
 - Math on distributions is well defined.



THE ROBOTICS INST



Continuous Random Variable

• Pdf – probability density function.



 Describes probability of each possible outcome of a single experiment.

Carnegie Mellon THE ROBOTICS INSTITUTE



THE ROBOTICS INS

(Joint) 2D Distributions







(Conditional) 2D Distributions



 $-\infty$

• Take a slice and renormalize.





$$p(x) = \frac{1}{\sqrt{2\pi\sigma}} exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)$$





N Dimensional Gaussian

• Formula:

$$p(\underline{x}) = \frac{1}{(2\pi)^{n/2} \sqrt{|C|}} exp\left(-\frac{[\underline{x}-\underline{\mu}]^T C^{-1} [\underline{x}-\underline{\mu}]}{2}\right)$$

• Mahalanobis distance:

$$\left[\underline{\mathbf{x}} - \underline{\boldsymbol{\mu}}\right]^{\mathrm{T}} \mathrm{C}^{-1} \left[\underline{\mathbf{x}} - \underline{\boldsymbol{\mu}}\right]$$

• C is "covariance matrix" defined later.

5.1.2.2 Expectation



Carnegie Mellon

THE ROBOTICS INSTIT

• For any function of x, this is just a weighted average where the pdf is the weight.

$$Exp[h(x)] = \int_{-\infty}^{\infty} h(x)p(x)dx \text{ scalar-scalar continuous}$$

$$Exp[\underline{h}(x)] = \int_{-\infty}^{\infty} \underline{h}(x)p(x)dx \text{ vector-scalar continuous} \text{ Notation means volume integral}$$

$$Exp[\underline{h}(x)] = \int_{-\infty}^{\infty} \underline{h}(\underline{x})p(\underline{x})d\underline{x} \text{ vector-vector continuous}$$

$$Exp[\underline{h}(x)] = \sum_{i=1}^{\infty} \underline{h}(\underline{u}_i)P(\underline{u}_i) \text{ vector-vector discrete}$$

• This is a functional or moment (with infinite limits of integration) so you need the entire pdf to work it out.



5.1.2.2 Expectation

• Properties inherited from integrals.

Exp[k] = k Exp[kh(x)] = kExp[h(x)]Exp[h(x) + g(x)] = Exp[h(x)] + Exp[g(x)]

Expectation is a linear operator over <u>functions</u>.





5.1.2.2 Mean

• Set $h(x) \rightarrow x$ etc.

$$\mu = Exp[x] = \int_{-\infty}^{\infty} [xp(x)]dx \qquad scalar \ continuous$$
$$\mu = Exp(\underline{x}) = \int_{-\infty}^{\infty} \underline{x}p(\underline{x})d\underline{x} \qquad vector \ continuous$$
$$\mu = Exp[\underline{x}] = \sum_{i=1}^{\infty} \underline{x}_i P(\underline{x}_i) \qquad vector \ discrete$$

• This is a property of the distribution of the population.



5.1.2.2 Mean and Most Likely Value





- Expected value is a centroid.
- It is not always the most likely value to occur.



Variance of a Random Scalar

• Set $h(x) \rightarrow [x-\mu]2$.

$$\sigma_{xx} = \int_{-\infty}^{\infty} \left[(x - \mu)^2 \cdot p(x) \right] dx$$

- Alternative notation: $\sigma^2(x)$
- Standard deviation defined as:

$$\sigma_{x} = \sigma(x) = \sqrt{\sigma_{xx}}$$





Recall : "Outer" Product

• Opposite of "inner" or dot product.

$$\underline{\mathbf{x}} = \begin{bmatrix} \mathbf{x} & \mathbf{y} & \mathbf{z} \end{bmatrix}^{\mathrm{T}}$$

• Generates a symmetric matrix from a vector.







Co-Variance of a Random Vector

• Continuous and discrete cases.

$$\Sigma = E([\underline{x} - \underline{\mu}][\underline{x} - \underline{\mu}]^{T}) = \int_{-\infty}^{\infty} [\underline{x} - \underline{\mu}][\underline{x} - \underline{\mu}]^{T}f(\underline{x})d\underline{x}$$

$$\Sigma = E([\underline{x} - \underline{\mu}][\underline{x} - \underline{\mu}]^{T}) = \sum_{i=1}^{n} [\underline{x} - \underline{\mu}][\underline{x} - \underline{\mu}]^{T}p(\underline{x})$$

Integral of a matrix is the matrix of the integrals.



• Elemental variances and covariances $s_{ii} = \frac{1}{n} \sum_{i=1}^{n} [x_i - \mu_i] [x_i - \mu_i] \quad s_{ij} = \frac{1}{n} \sum_{i=1}^{n} [x_i - \mu_i] [x_j - \mu_j]$

Carnegie Mel

THE ROBOTICS INST

5.1.2.3 Sampling Distributions and Statistics

• "Batch" Methods:

$$\mathbf{\bar{x}} = \frac{1}{n} \sum_{i=1}^{n} \mathbf{x}$$

$$S = \frac{1}{n} \sum_{i=1}^{n} [\underline{x} - \underline{\mu}] [\underline{x} - \underline{\mu}]^{T}$$

• Not feasible computationally for continuous update when N is large.

5.1.2.4 Computing Sample Statistics

• "Recursive" Methods:

$$\overline{\underline{x}}_{k+1} = \frac{(k\overline{\underline{x}}_k + \underline{x}_{k+1})}{(k+1)}$$
$$S_{k+1} = \frac{kS_k + [\underline{x}_{k+1} - \underline{\mu}][\underline{x}_{k+1} - \underline{\mu}]^T}{(k+1)}$$

• Related to the Kalman Filter.



Computing Sample Statistics

• "Calculator" Methods use accumulators:

$$\begin{array}{ll} \text{mean} & \underline{T}_{k+1} = \underline{T}_k + \underline{x}_{k+1} & \text{when data arrives} \\ & \underline{\bar{x}}_{k+1} = \frac{T_{k+1}}{(k+1)} & \text{when answer necessary} \\ \\ \text{covariance} & Q_{k+1} = Q_k + [\underline{x}_{k+1} - \underline{\mu}][\underline{x}_{k+1} - \underline{\mu}]_{\text{when data arrives}} \\ & S_{k+1} = \frac{Q_k}{(k+1)} & \text{when answer necessary} \end{array}$$

Carnegie Me

THE ROBOTICS INST

• Used in ... you guessed it ... hand calculators.

Contours of Constant Probability

• Consider the probability contained within a symmetric interval on the x axis.



Carnegie Mellon THE ROBOTICS INSTITUTE

5.1.2.6 Contours of Constant Probability

• In 2D, consider contours of constant exponent.



• These are ellipsoids in n dimensions:

$$(\underline{\mathbf{x}} - \underline{\boldsymbol{\mu}})^{\mathrm{T}} \Sigma^{-1} (\underline{\mathbf{x}} - \boldsymbol{\mu}) = k^2(p)$$



Outline

- 5.1 Random Variables, Processes and Transformation
 - 5.1.1 Characterizing Uncertainty
 - 5.1.2 Random Variables
 - 5.1.3 Transformation of Uncertainty
 - 5.1.4 Random Processes
 - Summary



Transformation

- "Pass covariance through a function":
- Suppose y = 2x and x is random.
 - 1st point: if x is random, y must be random even if "2" is not.
 - 2nd point: if we know cov(x) we can find cov(y). How?
 Here's the hard way.



THE ROBOT

Linearization

 The Taylor series allows us to extend any function into a neighborhood around a given point if we know the derivatives at that point:

$$f(x + \Delta x) = f(x) + f'(x)\Delta x + f''(x)\frac{\Delta x^2}{2} + \dots$$

- Error involved in truncation is related to magnitude of first neglected term.
- We linearize like so:

$$f(x + \Delta x) \approx f(x) + f'(x)\Delta x$$

• Errors involved are "second order"



"Analytic

5.1.3.1 Linear Transformation: Mean

• Suppose we know μ_x and want μ_y where:

$$\underline{y} = F\underline{x}$$
 F independent of x

 Because expectation is an integral and hence a linear operator:

$$\mu_{\underline{y}} = Exp(F\underline{x}) = FExp(\underline{x})$$

In other words

$$\mu_{\underline{y}} = F\mu_{\underline{x}}$$



5.1.3.1 Linear Transformation: Covariance

- Suppose we know σ_x and want σ_y where: y = Fx Findependent of x
- Because covariance is an integral and hence a linear operator:

$$\Sigma_{yy} = Exp(F\underline{x}\underline{x}^{T}F^{T}) = FExp(\underline{x}\underline{x}^{T})F^{T}$$

• In other words

$$\Sigma_{\underline{y}\underline{y}} = F\Sigma_{\underline{x}\underline{x}}F^T$$



5.1.3.2 Variance of a Sum of RVs

 Suppose there are n random variables x_i of same distribution.

$$x_i \sim N(\mu, \sigma)$$
, $i = 1, n$

Variance of x'es known and equal.

• Define a new variable y as the sum of these:

 What is the variance of y?





5.1.3.2 Variance of a Sum of RVs

• By our rules for uncertainty transforms:

$$\Sigma_{\underline{y}\underline{y}} = F\Sigma_{\underline{x}\underline{x}}F^{T}$$

- Where, in this case: $F = \begin{bmatrix} 1 & 1 & \dots & l \end{bmatrix}$
- Hence: $\sigma_y^2 = F \Sigma_{xx} F^T = \sum \sigma_{xy}^2$

• IOW:

$$\sigma_y^2 = F \Sigma_{\underline{x}\underline{x}} F^T = \sum_{i=1}^{\infty} \sigma_y^2 = n \sigma_x^2$$

Carnegie Mellon THE ROBOTICS INSTITUTE

n

5.1.3.3 Variance of an Average of RVs

 Suppose there are n random variables x_i of same distribution.

$$x_i \sim N(\mu, \sigma)$$
, $i = 1, n$

Variance of x'es known and equal.

• Define a new variable y as the average of these:



 What is the variance of y?



5.1.3.3 Variance of an Average of RVs

• By our rules for uncertainty transforms:

$$\Sigma_{\underline{y}\underline{y}} = F\Sigma_{\underline{x}\underline{x}}F^T$$

- Where, in this case: $F = \frac{l}{n} \begin{bmatrix} 1 & 1 & \dots & l \end{bmatrix}$
- Hence: $\sigma_y^2 = J \Sigma_{\underline{x}\underline{x}} J^T = \frac{1}{n^2} \sum \sigma_{x_i}^2$

• IOW:

$$\sigma_y^2 = \frac{1}{n}\sigma_x^2$$

Carnegie Mellon THE ROBOTICS INSTITUTE

n

i = 1

5.1.3.4 Coordinate Transformations



- Know covariance in one frame (because its easy to express there).
- Want to know it in another frame.



5.1.3.4 Coordinate Transformations



THE ROBOTICS IN

5.1.3.5 Nonlinear Transformation: Mean

- Suppose we know μ_x and want μ_y where: $\underline{y} = f(\underline{x})$
- Write x in terms of a deviation from a reference x': x = x' + ε
- Can use Jacobian to linearize: $y = f(x) = f(x' + \varepsilon) \approx f(x') + J\varepsilon$
- The mean of the distribution $\mu_{\underline{y}} = E_{\underline{y}}$ of y is....
- x' is not random, so...
- If e is unbiased, then......
- And if we choose Mean of the f() is the f() of the $\mu_{\underline{y}} = Exp(\underline{y}) = f(\mu_{\underline{x}})$

$$\mu_{\underline{y}} = Exp(\underline{y}) \approx Exp[f(\underline{x}')] + Exp(J\underline{\varepsilon})$$

$$Exp[f(\underline{x}')] = f(\underline{x}')$$

$$Exp(J\underline{\varepsilon}) = JExp(\underline{\varepsilon}) = 0$$

$$\underline{x}' = \mu_{\underline{x}}$$
"to first order"
"for unbiased error"

Carnegie Mellon THE ROBOTICS INSTITUTE

mean.

5.1.3.5 NonLinear Transformation: Covariance

- **Rewriting:**
- By definition:
- Which is:



$$\Sigma_{\underline{y}\underline{y}} = Exp([\underline{y} - \underline{y}'][\underline{y} - \underline{y}']^T)$$

$$\Sigma_{\underline{y}\underline{y}} = Exp(J\underline{\varepsilon}\underline{\varepsilon}^T J^T)$$

 $\Sigma_{yy} = J\Sigma$



Linearization: Again

• Whenever you write:

$$\Sigma_{\underline{y}\underline{y}} = J\Sigma_{\underline{x}\underline{x}}J^{T}$$

- Unless all derivatives beyond J vanish (i.e unless the mapping from x to y really is linear)
 - You have written an approximation.



5.1.3.6 Covariance with Partitioned Inputs

- Suppose we have: $\underline{y} = f(\underline{x}) \quad \underline{x} = \begin{bmatrix} \underline{x}_1 & \underline{x}_2 \end{bmatrix}^T$
- Partition the Jacobian and the Covariance:

$$\mathbf{J}_{\mathbf{x}} = \begin{bmatrix} \mathbf{J}_1 & \mathbf{J}_2 \end{bmatrix}$$

• We already know that the covariance of y is:

$$\Sigma_{yy} = J_x \Sigma_{xx} J_x^T \implies$$



Uncorrelated Partitioned Inputs

• Suppose we have:

$$\Sigma_{12} = \Sigma_{21} = [0]$$



$$\Sigma_{y} = J_{1}\Sigma_{11}J_{1}^{T} + J_{2}\Sigma_{22}J_{2}^{T}$$

Carnegie Mellon THE ROBOTICS INSTITUTE

Uncertainties of

Box 5.1 Formulae for Transformation of Uncertainty

For the following nonlinear transformation relating random vector \underline{x} to random vector \underline{y} :

$$\underline{y} = \underline{f}(\underline{x})$$

The mean and covariance of \underline{y} are related to those of \underline{x} by:

$$\mu_{\underline{y}} = \underline{f}(\mu_{\underline{x}}) \qquad \qquad \Sigma_{\underline{y}\underline{y}} = J\Sigma_{\underline{x}\underline{x}}J^{T}$$

Remember that, when using this result, unless all derivatives beyond J vanish (unless the original mapping really was linear), the result is a linear approximation to the true mean and covariance.

When x can be partitioned into two uncorrelated components, then:

$$\Sigma_{\underline{y}\underline{y}} = J_1 \Sigma_{11} J_1^T + J_2 \Sigma_{22} J_2^T$$

Carnegie Mellon THE ROBOTICS INSTITUTE

5.1.3.7 Example : Azimuth Scanner Transforming Uncertainty from 's' to 'w'



THE ROBOTICS INST

Step 1: From i to s



• Differentiate:

$$J_{i}^{s} = \begin{array}{c} \mathbf{R} & \mathbf{\psi} & \mathbf{\theta} \\ \mathbf{z} & \mathbf{x} \\ \mathbf{y}_{i}^{s} = \mathbf{y} \\ \mathbf{z} \\ -s\psi & -\mathbf{R}c\psi & \mathbf{0} \\ -c\psi s\theta & \mathbf{R}s\psi s\theta & -\mathbf{R}c\psi c\theta \end{array}$$



Carnegie Mellon THE ROBOTICS INSTITUTE

Step 1: Transformation



Step 2: From s to w

- T^w_smatrix relates s to w.
- Translation part is additive and irrelevant, so....

$$\Sigma_{\rm w} = R_{\rm s}^{\rm w} \Sigma_{\rm s} (R_{\rm s}^{\rm w})^{\rm T}$$

$$\Sigma_{w} = R_{s}^{w} J_{i}^{s} \Sigma_{i} (J_{i}^{s})^{T} (R_{s}^{w})^{T}$$



$$T_{s}^{w} = \operatorname{Trans}(0, 0, h)\operatorname{Roty}(\theta)$$
$$T_{s}^{w} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & h \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} c\theta & 0 & s\theta & 0 \\ 0 & 1 & 0 & 0 \\ -s\theta & 0 & c\theta & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

5.1.3.8 Example: Attitude from Terrain Map



- Find uncertainty in computed pitch angle given uncertainty in terrain
 - Which came from uncertainty in sensor.

5.1.3.8 Example: Attitude from Terrain Map

- Suppose the uncertainty in elevation is:
- Variance of computed pitch angle is:
- Where the Jacobian in this case is a gradient:
- The result is:

$$\sigma_{\theta}^2 = \frac{1}{L^2} [\sigma_f^2 + \sigma_r^2]$$

 $\Sigma_{z} = \begin{bmatrix} \sigma_{f}^{2} & o \\ 0 & \sigma_{r}^{2} \end{bmatrix}$

 $\Sigma_{A} = J \Sigma_{z} J^{T}$

 $\mathbf{J} = \begin{vmatrix} \frac{\partial \theta}{\partial z_{f}} & \frac{\partial \theta}{\partial z_{r}} \end{vmatrix} = \begin{bmatrix} \frac{1}{L} & -\frac{1}{L} \end{bmatrix}$



5.1.3.9 Example: Range Error in Rangefinders

- Where do the input variances come from?
- Variance in measured range depends on Range (R) reflectance (r) and incidence angle (a).



$$\sigma_{\rm R} \propto \left[\frac{\lambda {\rm R}^2}{\rho \cos \alpha}\right]$$



5.1.3.9 Example: Range Error in Rangefinders (Real Data)



RI THE ROBOTICS INSTITU

5.1.3.9 Example: Range Error in Rangefinders (Real Data)



THE ROBOTICS INST

5.1.3.10 Example: Stereo Vision

• From similar triangles:

$$\frac{Y_L}{X_L} = \frac{Y_L}{X} = \frac{y_l}{f} \quad \frac{Y_R}{X_R} = \frac{Y_R}{X} = \frac{y_r}{f}$$

• Subtract:

$$Y_{L} - Y_{R} = \frac{X[x_{l} - x_{r}]}{f}$$

• Hence:

$$X = \frac{bf}{d} \qquad b = \frac{Xd}{f}$$



Carnegie Mellon THE ROBOTICS INSTITUTE

62

5.1.3.11 Example: Stereo Uncertainty

- Define the normalized disparity:
- Now triangulation looks like:
- Uncertainty transformation:
- Jacobian is the scalar:
- Variance goes with 4th power of range:
- Standard deviation with the square of range.
 - Famous result.

$$X = \frac{b}{\delta}$$
$$\sigma_{xx} = J \sigma_{\delta\delta} J^{T}$$
$$\partial X = -h$$

 $\delta = \frac{d}{f} = \frac{b}{X}$

$$\mathbf{J} = \frac{\partial \mathbf{X}}{\partial \delta} = \frac{-\mathbf{b}}{\delta^2} \quad \blacksquare$$

$$\sigma_{xx} = \left[\frac{b^2}{\delta^4}\right] \sigma_{\delta\delta} = \left[\frac{X^4}{b^2}\right] \sigma_{\delta\delta}$$

THE ROBOTICS IN

Outline

- 5.1 Random Variables, Processes and Transformation
 - 5.1.1 Characterizing Uncertainty
 - 5.1.2 Random Variables
 - 5.1.3 Transformation of Uncertainty
 - 5.1.4 <u>Random Processes</u> SKIP
 - Summary



Outline

- 5.1 Random Variables, Processes and Transformation
 - 5.1.1 Characterizing Uncertainty
 - 5.1.2 Random Variables
 - 5.1.3 Transformation of Uncertainty
 - 5.1.4 Random Processes
 - <u>Summary</u>



Summary

- There are many kinds of error.
 - They can be removed with calibration, filtering.
- Covariance measures spread. Level curves are ellipsoids.
- Covariance is transformed with a matrix quadratic form.
- Variance of a random walk process grows linearly with time.
- Stochastic Diff Eqs are almost as easy to solve as deterministic ones.