Second-Moment Characterization of Random Variables and Vectors.

Second-Moment (SM) and First-Order Second-Moment (FOSM) Propagation of Uncertainty

(a) Random Variables

- **Second-Moment Characterization**

  - **Mean (expected value) of a random variable**
    
    \[ E[X] = m_X = \sum_{x_i} x_i P_X(x_i) \quad \text{(discrete case)} \]
    
    \[ = \int_{-\infty}^{\infty} x f_X(x) dx \quad \text{(continuous case)} \]

  - **Variance (second central moment) of a random variable**
    
    \[ \sigma_X^2 = Var[X] = E[(X - m_X)^2] = \sum_{x_i} (x_i - m_X)^2 P_X(x_i) \quad \text{(discrete case)} \]
    
    \[ \sigma_X^2 = \int_{-\infty}^{\infty} (x - m_X)^2 f_X(x) dx \quad \text{(continuous case)} \]

- **Examples**

  - **Poisson distribution**
    
    \[ P_Y(y) = \frac{(\lambda t)^y e^{-\lambda t}}{y!}, \quad y = 0, 1, 2, \ldots \]
    
    \[ m_Y = \lambda t \]
    
    \[ \sigma_Y^2 = \sum_{y=0}^{\infty} (y - \lambda t)^2 P_Y(y) = \lambda t = m_Y \]
- **Exponential distribution**

  \[ f_X(x) = \lambda e^{-\lambda x}, \quad x \geq 0 \]

  \[ m_X = \frac{1}{\lambda} \]

  \[ \sigma^2_X = \int_0^\infty \left( x - \frac{1}{\lambda} \right)^2 f_X(x) dx = \frac{1}{\lambda} = m_X^2 \]

- **Notation**

  \( X \sim (m, \sigma^2) \) indicates that \( X \) is a random variable with mean value \( m \) and variance \( \sigma^2 \).

- **Other measures of location**

  - **Mode** \( \tilde{x} \) = value that maximizes \( P_X \) or \( f_X \)
  
  - **Median** \( x_{50} \) = value such that \( F_X(x_{50}) = 0.5 \)

- **Other measures of dispersion**

  - **Standard deviation**

    \[ \sigma_X = \sqrt{\sigma^2_X} \] (same dimension as \( X \))

  - **Coefficient of variation**

    \[ V_X = \frac{\sigma_X}{m_X} \] (dimensionless quantity)

- **Expectation of a Function of a Random Variable. Initial and Central Moments.**

  - **Expected value of a function of a random variable**

    Let \( Y = g(X) \) be a function of a random variable \( X \). Then the mean value of \( Y \) is:
$E[Y] = E[g(X)] = \int_{-\infty}^{\infty} y f_Y(y) dy$

Importantly, it can be shown that $E[Y]$ can also be found directly from $f_X$, as:

$E[Y] = \int_{-\infty}^{\infty} g(x) f_X(x) dx$

- **Linearity of expectation**

It follows directly from the above and from linearity of integration that, for any constants $a_1$ and $a_2$ and for any functions $g_1(X)$ and $g_2(X)$:

$E[a_1 g_1(X) + a_2 g_2(X)] = a_1 E[g_1(X)] + a_2 E[g_2(X)]$

- **Expectation of some important functions**

1. $E[X^n] = \int_{-\infty}^{\infty} x^n f_X(x) dx$

   (called initial moments; the mean $m_X$ is also the first initial moment)

2. $E[(X - m_X)^n] = \int_{-\infty}^{\infty} (x - m_X)^n f_X(x) dx$

   (called central moments; the variance $\sigma_X^2$ is also called the second central moment)

- **Consequences of Linearity of Expectation. Second-Moment(SM) Propagation of Uncertainty for Linear Functions.**

1. $\sigma_X^2 = Var[X] = E[(X - m_X)^2] = E[X^2] - 2m_X E[X] + m_X^2 = E[X^2] - m_X^2$

   $\Rightarrow E[X^2] = \sigma_X^2 + m_X^2$

2. Let $Y = a + bX$, where $a$ and $b$ are constants. Using linearity of expectation, one obtains the following expressions for the mean value and variance of $Y$:

   $m_Y = a + b E[X] = a + bm_X$

   $\sigma_Y^2 = E[(Y - m_Y)^2] = b^2 \sigma_X^2$
• **First-Order Second-Moment (FOSM) Propagation of Uncertainty for Nonlinear Functions**

Usually, with knowledge of only the mean value and variance of $X$, it is impossible to calculate $m_Y$ and $\sigma_Y^2$. However, a so-called first-order second-moment (FOSM) approximation can be obtained as follows.

Given $X \sim (m_X, \sigma_X^2)$ and $Y = g(X)$, a generic nonlinear function of $X$, find the mean value and variance of $Y$.

Replace $g(X)$ by a linear function of $X$, usually by linear Taylor expansion around $m_X$. This gives the following approximation to $g(X)$:

$$Y = g(X) \approx g(m_X) + \left. \frac{dg(X)}{dX} \right|_{m_X} (X - m_X)$$

Then approximate values for $m_Y$ and $\sigma_Y^2$ are:

$$m_Y = g(m_X), \quad \sigma_Y^2 = \left( \left. \frac{dg(X)}{dX} \right|_{m_X} \right)^2 \sigma_X^2$$

(b) **Random Vectors**

• **Second-Moment Characterization. Initial and Central Moments.**

Consider a random vector $\mathbf{X}$ with components $X_1, X_2, \ldots, X_n$.

- **Expected value**

$$E[\mathbf{X}] = E \begin{bmatrix} X_1 \\ \vdots \\ X_n \end{bmatrix} = \begin{bmatrix} E[X_1] \\ \vdots \\ E[X_n] \end{bmatrix} = \begin{bmatrix} m_1 \\ \vdots \\ m_n \end{bmatrix} = \mathbf{m}$$

(mean value vector)

- **Expected value of a scalar function of $\mathbf{X}$**

Let $Y = g(\mathbf{X})$ be a function of $\mathbf{X}$. Then, extending a result given previously for functions of single variables, one finds that $E[Y]$ may be calculated as:
\[ E[Y] = \int_{\mathbb{R}^n} g(x) f_X(x) dx \]

Again, it is clear that linearity applies, in the sense that, for any given constants \( a_1 \) and \( a_2 \) and any given functions \( g_1(X) \) and \( g_2(X) \):

\[ E[a_1 g_1(X) + a_2 g_2(X)] = a_1 E[g_1(X)] + a_2 E[g_2(X)] \]

- **Expectation of some special functions**
  
  - **Initial moments**
    
    1. Order 1: \( E[X_i] = m_i \Leftrightarrow E[X] = m, \quad i = 1, 2, \ldots, n \)
    
    2. Order 2: \( E[X_i X_j] = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x_i x_j f_{X_i, X_j}(x_i, x_j) dx_i dx_j, \quad i, j = 1, 2, \ldots, n \)
    
    3. Order 3: \( E[X_i X_j X_k] = \ldots, \quad i, j, k = 1, 2, \ldots, n \)

  
  - **Central moments**
    
    1. Order 1: \( E[X_i - m_i] = 0, \quad i = 1, 2, \ldots, n \)
    
    2. Order 2 (covariance between two variables):

    \[
    \text{Cov}[X_i, X_j] = E[(X_i - m_i)(X_j - m_j)], \quad i, j = 1, 2, \ldots, n
    \]
    
    \[
    = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (x_i - m_i)(x_j - m_j) f_{X_i, X_j}(x_i, x_j) dx_i dx_j
    \]

  
  - **Covariance in terms of first and second initial moments**

    Using linearity of expectation,

    \[
    \text{Cov}[X_i, X_j] = E[(X_i - m_i)(X_j - m_j)] = E[X_i X_j - X_i m_j - m_i X_j + m_i m_j]
    = E[X_i X_j] - m_i m_j
    \Rightarrow E[X_i X_j] = \text{Cov}[X_i, X_j] + m_i m_j
    \]
\textbf{Covariance Matrix and Correlation Coefficients}

- \textit{Covariance matrix}

\[
\Sigma_{X} = \begin{bmatrix}
\text{Cov}[X_i, X_j] \\
\cdot \\
\cdot \\
\cdot \\
(i, j = 1, 2, \ldots, n)
\end{bmatrix}
\]
\[
= E[(X - m_x)(X - m_x)^T]
\]

- For \( n = 2 \):

\[
\Sigma_X = \begin{bmatrix}
\sigma_1^2 & \text{Cov}[X_1, X_2] \\
\text{Cov}[X_2, X_1] & \sigma_2^2
\end{bmatrix}
\]

- \( \Sigma_X \) is the matrix equivalent of \( \sigma_X^2 \)

- \( \Sigma_X \) is symmetrical: \( \Sigma_X = \Sigma_X^T \)

- \textit{Correlation coefficient between two variables}

\[
\rho_{ij} = \frac{\text{Cov}[X_i, X_j]}{\sigma_i \sigma_j}, \quad i, j = 1, 2, \ldots, n, \quad -1 \leq \rho_{ij} \leq 1
\]

- \( \rho_{ij} \) is a measure of linear dependence between two random variables;

- \( \rho_{ij} \) has values between -1 and 1, and is dimensionless.
• SM Propagation of Uncertainty for Linear Functions of Several Variables

Let $Y = a_0 + \sum_{i=1}^{n} a_i X_i = a_0 + a_1 X_1 + a_2 X_2 + \cdots + a_n X_n$ be a linear function of the vector $X$. Using linearity of expectation, one finds the following important results:

\[
E[Y] = E \left[ a_0 + \sum_{i=1}^{n} a_i X_i \right] = a_0 + \sum_{i=1}^{n} a_i m_i
\]

\[
Var[Y] = \sum_{i=1}^{n} a_i^2 Var[X_i] + 2 \sum_{i=1}^{n} \sum_{j=i+1}^{n} a_i a_j Cov[X_i, X_j]
\]
• For \( n = 2 \):

\[
Y = a_0 + a_1 X_1 + a_2 X_2
\]

\[
E[Y] = a_0 + a_1 E[X_1] + a_2 E[X_2]
\]

\[
Var[Y] = a_1^2 Var[X_1] + a_2^2 Var[X_2] + 2a_1a_2 Cov[X_1, X_2]
\]

• For uncorrelated random variables:

\[
Var[Y] = \sum_{i=1}^{n} a_i^2 Var[X_i]
\]

• Extension to several linear functions of several variables

Let \( Y \) be a vector whose components \( Y_i \) are linear functions of a random vector \( X \). Then, one can write \( Y = a + B X \), where \( a \) is a given vector and \( B \) is a given matrix. One can show that:

\[
m_Y = a + B m_X
\]

\[
\Sigma_Y = B \Sigma_X B^T
\]

• FOSM Propagation of Uncertainty for Nonlinear Functions of Several Variables

Let \( X \sim (m_X, \Sigma_X) \) be a random vector with mean value vector \( m_X \) and covariance matrix \( \Sigma_X \). Consider a nonlinear function of \( X \), say \( Y = g(X) \). In general, \( m_Y \) and \( \sigma_Y^2 \) depend on the entire joint distribution of the vector \( X \). However, simple approximations to \( m_Y \) and \( \sigma_Y^2 \) are obtained by linearizing \( g(X) \) and then using the exact SM results for linear functions. If linearization is obtained through linear Taylor expansion about \( m_X \), then the function that replaces \( g(X) \) is:

\[
g(X) \approx g(m_X) + \sum_{i=1}^{n} \frac{\partial g(X)}{\partial X_i} \bigg|_{X=m_X} (X_i - m_i)
\]

where \( m_i \) is the mean value of \( X_i \). The approximate mean and variance of \( Y \) are then:

\[
m_Y = g(m_X),
\]

\[
\sigma_Y^2 = \sum_{i=1}^{n} \frac{\partial g(X)}{\partial X_i} \bigg|_{X=m_X} \Sigma_X \bigg|_{X=m_X} \frac{\partial g(X)}{\partial X_i} \bigg|_{X=m_X}^T
\]
\[ \sigma_Y^2 = \sum_{i=1}^{n} \sum_{j=1}^{n} b_i b_j \text{Cov}[X_i, X_j] \]

where

\[ b_i = \left. \frac{\partial g(X)}{\partial X_i} \right|_{X=m_X} \]

This way of propagating uncertainty is called FOSM analysis.