1 Stochastic Processes

1.1 Deterministic Linear Systems

Continuous linear systems are often described by a set of n differential equations

$$\dot{\boldsymbol{x}} = A(t)\boldsymbol{x} + B(t)\boldsymbol{u}, \quad \boldsymbol{x}_0 = given \tag{1}$$

$$\mathbf{y} = H(t)\boldsymbol{x} \tag{2}$$

where $\boldsymbol{x}$ is an n-dimensional state vector, $\boldsymbol{u}$ is an m-dimensional control input, $\mathbf{y}$ is an l-dimensional output or measurement, $A(t)$ is $n \times n$ matrix $B(t)$ is $n \times m$ matrix, and $C(t)$ is an $l \times n$ matrix. The solution to these linear differential equations is

$$\boldsymbol{x}(t) = \phi(t, t_0)\boldsymbol{x}_0 + \int_0^t \phi(t, \tau)B(\tau)u(\tau)d\tau \tag{3}$$

where the state transition matrix, $\phi(t, \tau)$, is the solution to the matrix differential equation

$$\dot{\phi}(t, t_o) = A(t)\phi(t, t_0), \quad \phi(t_0, t_0) = I \tag{4}$$

When the system is time-invariant (i.e. when A, B, H, are constant), the solution to this differential equation is simple

$$\phi(t, t_0) = \phi(t - t_0) = \exp\{A(t - t_0)\} \tag{5}$$

where $\exp\{A(t - t_0)\}$ is the matrix exponential function,

$$\phi(t - t_0) = \exp\{A(t - t_0)\} = I + A(t - t_0) + A^2(t - t_0)^2/2! + A^3(t - t_0)^3/3! + \cdots \tag{6}$$

1.2 The Expectation (Averaging) Operator

The expectation operation, $E_x[\cdot]$, determines the average value of a function with respect to a random variable x given the probability density function $f_x(\xi)$.

$$E_x[\cdot] = \int_{-\infty}^{+\infty} (\cdot)\, f_x(\xi)d\xi \tag{7}$$

The expected value, mean, or average value of x is

$$\bar{x} = E_x[x] = \int_{-\infty}^{+\infty} \xi\, f_x(\xi)d\xi \tag{8}$$

The variance of x is defined as

$$\sigma_x^2 = E[(x - \bar{x})^2] = \int_{-\infty}^{+\infty} (\xi - \bar{x})^2\, f_x(\xi)d\xi \tag{9}$$

Notice that the expected value of a constant k is equal to k.

$$E_x[k] = \int_{-\infty}^{+\infty} k\, f_x(\xi)d\xi = k\int_{-\infty}^{+\infty} f(\xi)d\xi = k \tag{10}$$

For multi-variable functions we have

$$E_{\boldsymbol{x}}[\cdot] = \int_{-\infty}^{+\infty} (\cdot)\, f_{\boldsymbol{x}}(\boldsymbol{\xi}) d\boldsymbol{\xi} \tag{11}$$

where $f_{\boldsymbol{x}}(\boldsymbol{\xi})$ is the joint probability density function. For example, the mean of the vector $\boldsymbol{x}$ is

$$\bar{\boldsymbol{x}} = E_{\mathbf{x}}[\boldsymbol{x}] = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \cdots \int_{-\infty}^{+\infty} \xi\, f_{\boldsymbol{x}}(\boldsymbol{\xi}) d\boldsymbol{\xi} \tag{12}$$

and the covariance of $\boldsymbol{x}$ is define as

$$P_{\boldsymbol{xx}} = E[(\boldsymbol{x} - \bar{\boldsymbol{x}})(\boldsymbol{x} - \bar{\boldsymbol{x}})^T] = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \cdots \int_{-\infty}^{+\infty} (\boldsymbol{\xi} - \bar{\boldsymbol{x}})(\boldsymbol{\xi} - \bar{\boldsymbol{x}})^T\, f_{\boldsymbol{x}}(\boldsymbol{\xi}) d\boldsymbol{\xi} \tag{13}$$

Two important properties of the expectation operator:

- The expectation of a sum of random variables is equal to the sum of the expectations
- The expectation of a constant is that constant

Examples:

$$E[x + y] = E[x] + E[y] = M_x + M_y \tag{14}$$

$$E[k] = k \tag{15}$$

$$E[k_1 x + k_2 y] = k_1 E[x] + k_2 E[y] = k_1 M_x + k_2 M_y \tag{16}$$

$$E[(x - M_x)^2] = E[x^2 - 2M_x x + M_x^2] = E[x^2] - 2M_x E[x] + M_x^2 = M_{x^2} - 2M_x^2 + M_x^2 = M_{x^2} - M_x^2 \tag{17}$$

1.3 Gaussian Random Variables

The pdf for a Gaussian random variable is given by

$$f_x(\xi) = \frac{1}{\sqrt{2\pi\sigma_x^2}} exp\left\{-\frac{1}{2}(\xi - M_x)^2\right\} \tag{18}$$

where the mean and variance of the Gaussian random variable are

$$M_x = E_x[x] = \int_{-\infty}^{+\infty} \xi f_x(\xi) d\xi \tag{19}$$

$$\sigma_x^2 = E_x[(x - M_x)^2] = \int_{-\infty}^{+\infty} (\xi - M_x)^2 f_x(\xi) d\xi \tag{20}$$

We often say that the random variable x, is Gaussian or "normal" with mean M_x,and variance σ_x^2, $x \sim N(M_x, \sigma_x^2)$.

When the random variable is an n-dimensional vector, the joint pdf for a Gaussian random variable is given by

$$f_{\boldsymbol{x}}(\boldsymbol{\xi}) = \frac{1}{(2\pi)^{n/2}\left|P_{\boldsymbol{xx}}\right|^{1/2}} \exp\left\{-\frac{1}{2}(\boldsymbol{\xi} - \boldsymbol{M_x})^T P_{\boldsymbol{xx}}^{-1}(\boldsymbol{\xi} - \boldsymbol{M_x})\right\} \tag{21}$$

where

$$\boldsymbol{M_x} = E_{\boldsymbol{x}}[\boldsymbol{x}] = \begin{bmatrix} M_{x_1} \\ M_{x_2} \\ \vdots \\ M_{x_n} \end{bmatrix} \tag{22}$$

$$P_{\boldsymbol{xx}} = E_{\boldsymbol{x}}[(\boldsymbol{x} - \boldsymbol{M_x})(\boldsymbol{x} - \boldsymbol{M_x})^T] = \begin{bmatrix} \sigma_{x_1}^2 & \rho_{x_1x_2}\sigma_{x_1}\sigma_{x_2} & \cdots & \rho_{x_1x_n}\sigma_{x_1}\sigma_{x_n} \\ \rho_{x_1x_2}\sigma_{x_1}\sigma_{x_2} & \sigma_{x_2}^2 & \cdots & \rho_{x_2x_n}\sigma_{x_2}\sigma_{x_n} \\ \vdots & \vdots & \ddots & \vdots \\ \rho_{x_1x_n}\sigma_{x_1}\sigma_{x_n} & \rho_{x_2x_n}\sigma_{x_2}\sigma_{x_n} & \cdots & \sigma_{x_n}^2 \end{bmatrix} \tag{23}$$

The *covariance* matrix P_{xx} is an $n \times n$ symmetric matrix, and $\rho_{x_ix_j}$is the correlation coefficient for states x_i and x_j. The correlation coefficient is always between -1 and $+1$. If two states are uncorrelated, $\rho_{x_ix_j} = 0$.

Functions of Gaussian Random Variables

If a random variable $\boldsymbol{y} = A\boldsymbol{x}$ is a linear function of an n-dimensional Gaussian random variable $\boldsymbol{x} \backsim N(\boldsymbol{M_x}, P_{\boldsymbol{xx}})$ where A is an $m \times n$ matrix, then $\boldsymbol{y}$ is also a Gaussian random variable, $\boldsymbol{y} \backsim N(\boldsymbol{M_y}, P_{\boldsymbol{yy}})$ with mean and covariance given by

$$\boldsymbol{M_y} = E_{\boldsymbol{y}}[\boldsymbol{y}] = E_{\boldsymbol{x}}[A\boldsymbol{x}] = AE_{\boldsymbol{x}}[\boldsymbol{x}] = A\boldsymbol{M}_x \tag{24}$$

$$P_{\boldsymbol{yy}} = E_{\boldsymbol{y}}[(\boldsymbol{y} - \boldsymbol{M_y})(\boldsymbol{y} - \boldsymbol{M_y})^T] = E_{\boldsymbol{x}}[A(\boldsymbol{x} - \boldsymbol{M_x})(\boldsymbol{x} - \boldsymbol{M_x})^T A^T] = AE_{\boldsymbol{x}}[(\boldsymbol{x} - \boldsymbol{M_x})(\boldsymbol{x} - \boldsymbol{M_x})^T]A^{T=}AP_{\boldsymbol{xx}}A^T \tag{25}$$

If $\boldsymbol{y} = \boldsymbol{f}(\boldsymbol{x})$ is a nonlinear function of a Gaussian random variable $\boldsymbol{x} \backsim N(\boldsymbol{M_x}, P_{\boldsymbol{xx}})$, the mean and variance of $\boldsymbol{y}$ are approximately given by

$$\boldsymbol{M}_y \approx \boldsymbol{f}(\boldsymbol{M_x}) \tag{26}$$

$$P_{\boldsymbol{yy}} \approx FP_{\boldsymbol{xx}}F^T, \quad F = \left.\frac{\partial \boldsymbol{f}}{\partial \boldsymbol{x}}\right|_{\boldsymbol{M_x}} \tag{27}$$

1.4 Nonlinear Stochastic Processes

The most simple stochastic process to consider is a discrete-time zero-mean white noise Gaussian process, $\boldsymbol{w}_{d,i}$, where $\boldsymbol{w}_{d,i}$ is an m-dimensional vector with mean and covariance defined as

$$\begin{aligned} E[\boldsymbol{w}_{d,i}] &= 0 \\ E[\boldsymbol{w}_{d,i}\boldsymbol{w}_{d,j}^T] &= Q_{d_i}\delta_{ij}, \end{aligned} \tag{28}$$

where δ_{ij} is a Kronecker delta function. At any time t_i, this process has a zero mean, covariance Q_{d_i}, and is uncorrelated in time.

A much more general stochastic process, a discrete-time nonlinear stochastic process, can be described by the following difference equation,

$$\boldsymbol{x}_{i+1} = \boldsymbol{h}(\boldsymbol{x}_i, \boldsymbol{w}_{d,i}), \tag{29}$$

where $\boldsymbol{x}_i$ is an n-dimensional state vector. The stochastic nature of this nonlinear system is due to either the initial conditions, the input $\boldsymbol{w}_{d,i}$, or both. At each instant in time the state $\boldsymbol{x}_i$ is a random variable with mean and covariance (first and second moments) denoted by $\bar{\boldsymbol{x}}_i$, and P_i,respectively.

A similar development can be made for continuous systems (though not rigorously) by introducing a continuous zero-mean white noise Gaussian process $\boldsymbol{w}(t)$, where $\boldsymbol{w}(t)$ is an m-dimensional vector with mean and covariance defined as

$$\begin{aligned} E[\boldsymbol{w}(t)] &= 0 \\ E[\boldsymbol{w}(t)\boldsymbol{w}(t+\tau)^T] &= Q(t)\delta(\tau), \end{aligned} \tag{30}$$

where $Q(t)$ is the "strength" of the white noise process. The corresponding continuous non-linear stochastic process can be be described by the following differential equation

$$\dot{\boldsymbol{x}} = \boldsymbol{f}(\boldsymbol{x}, \boldsymbol{w}), \tag{31}$$

Once again, the stochastic nature of this non-system is due to either the initial conditions, the input $\boldsymbol{w}$, or both. At each instant in time, $\boldsymbol{x}(t)$, is a random variable with mean and covariance denoted by $\bar{\boldsymbol{x}}(t)$, and $P(t)$, respectively.

1.5 Linear Stochastic Processes

1.5.1 Discrete Linear Stochastic Processes

A discrete linear Gaussian stochastic process can be described by the n difference equations

$$\boldsymbol{x}_{i+1} = \Phi_i \boldsymbol{x}_i + \Gamma_i \boldsymbol{w}_{d,i}, \quad \boldsymbol{x}_0 \sim N(\bar{\boldsymbol{x}}_0,\, P_0) \tag{32}$$

where Φ_i is $n \times n$ matrix, Γ_i is $n \times m$ matrix, and $\boldsymbol{w}_{d,i}$ is an m-dimensional white noise process,

$$\begin{aligned} E[\boldsymbol{w}_{d,i}] &= 0 \\ E[\boldsymbol{w}_{d,i}\boldsymbol{w}_{d,j}^T] &= Q_{d_i}\delta_{ij} \end{aligned} \tag{33}$$

where Q_{d_i} , the covariance of the white noise, is an $m \times m$ matrix.

1.5.2 Continuous Linear Stochastic Processes

A continuous linear Gaussian stochastic process can be described by

$$\dot{\boldsymbol{x}} = A(t)\boldsymbol{x} + B(t)\boldsymbol{w}, \quad \boldsymbol{x}_0 \sim N(\bar{\boldsymbol{x}}_0,\, P_0) \tag{34}$$

where $A(t)$ is $n \times n$ matrix $B(t)$ is $n \times m$ matrix, and where $\boldsymbol{w}$ is an m-dimensional white noise process,

$$\begin{aligned} E[\boldsymbol{w}(t)] &= 0 \\ E[\boldsymbol{w}(t)\boldsymbol{w}(t')^T] &= Q(t)\delta(t-t') \end{aligned} \tag{35}$$

where $Q(t)$, the strength of the white noise, is an $m \times m$ matrix.

1.6 Time Propagation of the Mean and Covariance of a Random Process

1.6.1 Discrete Linear Systems

Consider a discrete linear Gaussian stochastic process described by the n difference equations

$$\boldsymbol{x}_{i+1} = \Phi_i \boldsymbol{x}_i + \Gamma_i \boldsymbol{w}_{d,i}, \quad \mathbf{x}_0 \sim N(\bar{\boldsymbol{x}}_0,\, P_0) \tag{36}$$

where Φ_i is $n \times n$ matrix, Γ_i is $n \times m$ matrix, and $\boldsymbol{w}_{d,i}$ is an m-dimensional white noise process,

$$\begin{aligned} E[\boldsymbol{w}_{d,i}] &= 0 \\ E[\boldsymbol{w}_{d,i}\boldsymbol{w}_{d,j}^T] &= Q_{d_i}\delta_{ij} \end{aligned} \tag{37}$$

where Q_{d_i} , the covariance of the white noise, is an $m \times m$ matrix.

The expected value or mean value of the state $\boldsymbol{x}_{i+1}$of this discrete linear stochastic process is given by

$$\boxed{\bar{\boldsymbol{x}}_{i+1} = \Phi_i \bar{\boldsymbol{x}}_i, \quad \bar{\boldsymbol{x}}_0 = given} \tag{38}$$

and the covariance of $\boldsymbol{x}_{i+1}$, $P_{i+1} = E[(\boldsymbol{x}_{i+1} - \bar{\boldsymbol{x}}_{i+1})(\boldsymbol{x}_{i+1} - \bar{\boldsymbol{x}}_{i+1})^T]$, is given by

$$\boxed{P_{i+1} = \Phi_i P_i \Phi_i^T + \Gamma_i Q_{d_i} \Gamma_i^T, \quad P_0 = given} \tag{39}$$

1.6.2 Continuous Linear Systems

Consider a continuous linear stochastic process described by

$$\dot{\boldsymbol{x}} = A(t)\boldsymbol{x} + B(t)\boldsymbol{w}, \quad \boldsymbol{x}_0 \sim N(\bar{\boldsymbol{x}}_0,\, P_0) \tag{40}$$

where $A(t)$ is $n \times n$ matrix $B(t)$ is $n \times m$ matrix, and where $\boldsymbol{w}$ is an m-dimensional white noise process,

$$\begin{aligned} E[\boldsymbol{w}(t)] &= 0 \\ E[\boldsymbol{w}(t)\boldsymbol{w}(t')^T] &= Q(t)\delta(t-t') \end{aligned} \tag{41}$$

where $Q(t)$, the strength of the white noise, is an $m \times m$ matrix.

The expected value or mean value of the stat e$\boldsymbol{x}(t)$of the continuous linear stochastic process is given by

$$\boxed{\dot{\bar{\mathbf{x}}} = A(t)\bar{\boldsymbol{x}} \quad \bar{\boldsymbol{x}}_0 = given} \tag{42}$$

and the covariance of $\boldsymbol{x}(t)$, $P(t) = E[(\boldsymbol{x}(t) - \bar{\boldsymbol{x}}(t))(\boldsymbol{x}(t) - \bar{\boldsymbol{x}}(t))^T]$, is given by

$$\boxed{\dot{P} = A(t)P(t) + P(t)A^T(t) + B(t)Q(t)B(t)^T, \quad P_0 = given} \tag{43}$$

Alternatively, if the linear system is discretized into time increments Δt such that,

$$\boldsymbol{x}_{i+1} = \Phi_i \boldsymbol{x}_i + \boldsymbol{w}_{d,i} \tag{44}$$

where

$$\Phi_i \approx e^{A(t_i)\Delta t} \tag{45}$$

$$\boldsymbol{w}_{d,i} \approx \int_t^{t+\Delta t} \Phi(t-\tau)B(\tau)w(\tau)d\tau \tag{46}$$

then the mean and covariance of the state can also be written as

$$\boxed{\bar{\boldsymbol{x}}_{i+1} = \Phi_i \bar{\boldsymbol{x}}_i, \quad \bar{\boldsymbol{x}}_0 = given} \tag{47}$$

$$\boxed{P_{i+1} = \Phi_i P_i \Phi_i^T + B_i Q B_i^T \Delta t, \quad P_0 = given} \tag{48}$$

1.6.3 Discrete Non-Linear Systems

Consider a discrete-time nonlinear stochastic process described by the following difference equation,

$$\boldsymbol{x}_{i+1} = \boldsymbol{h}(\boldsymbol{x}_i, \boldsymbol{w}_{d,i}), \quad \boldsymbol{x}_0 \sim N(\bar{\boldsymbol{x}}_0,\, P_0) \tag{49}$$

where $\boldsymbol{x}_i$ is an n-dimensional state vector, and where $\boldsymbol{w}_{d,i}$ is an m-dimensional vector with mean and covariance defined as

$$\begin{aligned} E[\boldsymbol{w}_{d,i}] &= 0 \\ E[\boldsymbol{w}_{d,i}\boldsymbol{w}_{d,j}^T] &= Q_{d_i}\delta_{ij}, \end{aligned} \tag{50}$$

Under the appropriate conditions (see derivation below), this non-linear difference equation can be linearized and the expected value and covariance of the state can be determined from linear stochastic system theory.

$$\boxed{\bar{\boldsymbol{x}}_{i+1} = \boldsymbol{h}(\bar{\boldsymbol{x}}_i, \boldsymbol{w}_{d,i}), \quad \bar{\boldsymbol{x}}_0 = given} \tag{51}$$

$$\boxed{P_{i+1} = H_i P_i H_i^T + \Gamma_i Q_{d_i} \Gamma_i^T, \quad P_0 = given} \tag{52}$$

where

$$H_i = \left.\frac{\partial \boldsymbol{h}_i(\boldsymbol{x}_i, \boldsymbol{w}_{d,i})}{\partial \boldsymbol{x}_i}\right|_{\bar{\boldsymbol{x}}, \bar{\boldsymbol{w}}_d}, \quad \Gamma_i = \left.\frac{\partial \boldsymbol{h}_i(\boldsymbol{x}_i, \boldsymbol{w}_{d,i})}{\partial \boldsymbol{w}_{d,i}}\right|_{\bar{\boldsymbol{x}}, \bar{\boldsymbol{w}}_d} \tag{53}$$

1.6.4 Continuous Non-Linear Systems

Consider a continuous non-linear stochastic process described by the following differential equation

$$\dot{\boldsymbol{x}} = \boldsymbol{f}(\boldsymbol{x}, \boldsymbol{w}), \quad \boldsymbol{x}_0 \sim N(\bar{\boldsymbol{x}}_0,\, P_0) \tag{54}$$

where $\boldsymbol{w}(t)$ is an m-dimensional vector with mean and covariance defined as

$$\begin{aligned} E[\boldsymbol{w}(t)] &= 0 \\ E[\boldsymbol{w}(t)\boldsymbol{w}(t+\tau)^T] &= Q(t)\delta(\tau), \end{aligned} \tag{55}$$

and where $Q(t)$ is the "strength" of the white noise process.

Under the appropriate conditions (see derivation below), this non-linear differential equation can be linearized and the the expected value and covariance of the state $\boldsymbol{x}(t)$ can be determined from linear stochastic system theory.

$$\boxed{\dot{\bar{x}} = \boldsymbol{f}(\bar{\boldsymbol{x}}, \bar{\boldsymbol{w}}), \quad \bar{\boldsymbol{x}}_0 = given} \tag{56}$$

$$\boxed{\dot{P} = F(t)P(t) + P(t)F^T(t) + B(t)Q(t)B(t)^T, \quad P_0 = given} \tag{57}$$

where

$$F(t) = \left.\frac{\partial \boldsymbol{f}(\boldsymbol{x},\boldsymbol{w})}{\partial \boldsymbol{x}}\right|_{\bar{\boldsymbol{x}}(t),\bar{\boldsymbol{w}}(t)}, \quad B(t) = \left.\frac{\partial \boldsymbol{f}(\boldsymbol{x},\boldsymbol{w})}{\partial \boldsymbol{w}}\right|_{\bar{\boldsymbol{x}}(t),\bar{\boldsymbol{w}}(t)} \tag{58}$$

Alternately, if the linearized equations are discretized, we can write

$$\boxed{P_{i+1} = \Phi_i P_i \Phi_i^T + B_i Q B_i^T \Delta t, \quad P_0 = given} \tag{59}$$

where

$$\Phi_i \approx e^{F(t_i)\Delta t} \tag{60}$$

1.7 Common Linear Stochastic Processes

1.7.1 A Scalar Discrete Random Walk

Suppose we have a scalar discrete linear system given by

$$x_{i+1} = x_i + w_{d,i} \tag{61}$$

where w_i is a discrete white noise process:

$$\begin{aligned} E[w_{d,i}] &= 0 \\ E[w_{d,i} w_{d,j}^T] &= Q_{d_i} \delta_{ij}, \end{aligned} \tag{62}$$

and where $Q_{d_i}$is typically an $n \times n$ diagonal matrix. What are the units of P_0? What are the units of Q_{d_i}? The mean or expected value of x_{i+1} is

$$\boxed{\bar{x}_{i+1} = \bar{x}_i, \quad \bar{x}_0 = given} \tag{63}$$

and the covariance is

$$\boxed{P_{i+1} = P_i + Q_{d_i}, \quad P_0 = given} \tag{64}$$

Note that $\bar{x}_{k+1} = \bar{x}_0$, and $P_{k+1} = P_0 + kQ_{d_i}$, i.e. the mean is constant and equal to the initial value, and the variance grows linearly in time.

1.7.2 A Scalar Continuous Random Walk

Suppose we have a scalar continuous linear system given by

$$\dot{x} = w \tag{65}$$

where w is a *continuous* white noise process:

$$\begin{aligned} E[w(t)] &= 0 \\ E[w(t)w(t')^T] &= Q\delta(t-t'), \end{aligned} \tag{66}$$

and where Q is typically $n \times n$ diagonal matrix. What are the units of P_0? What are the units of Q? What is the relationship between Q and Q_d?

The mean or expected value of x is propagated using

$$\boxed{\dot{\bar{x}} = 0, \quad \bar{x}_0 = given} \tag{67}$$

and the covariance propagated using

$$\boxed{\dot{P} = Q, \quad P_0 = given} \tag{68}$$

Note that $\bar{x}(t) = \bar{x}_0$, and $P(t) = P_0 + Qt$, i.e. the mean is constant and equal to the initial value, and the variance grows linearly in time.

1.7.3 A Scalar Discrete Random Bias

Suppose we have a scalar discrete linear system given by

$$x_{i+1} = x_i \tag{69}$$

The mean or expected value of x_{i+1} is

$$\boxed{\bar{x}_{i+1} = \bar{x}_i, \quad \bar{x}_0 = given} \tag{70}$$

and the covariance is

$$\boxed{P_{i+1} = P_i, \quad P_0 = given} \tag{71}$$

Note that $\bar{x}_{n+1} = \bar{x}_0$, and $P_{n+1} = P_0$, i.e. the mean and variance are constant.

1.7.4 A Scalar Continuous Random Bias

Suppose we have a scalar continuous linear system given by

$$\dot{x} = 0 \tag{72}$$

The mean or expected value of x is propagated using

$$\boxed{\dot{\bar{x}} = 0, \quad \bar{x}_0 = given} \tag{73}$$

and the covariance propagated using

$$\boxed{\dot{P} = 0, \quad P_0 = given} \tag{74}$$

Note that $\bar{x}(t) = \bar{x}_0$, and $P(t) = P_0$, i.e. the mean and variance are constant.

1.7.5 A Scalar Discrete 1st-Order Stochastic Process

Suppose we have a scalar discrete linear system described by the *scalar* equation

$$x_{i+1} = Kx_i + w_{d,i} \tag{75}$$

where w_i is a discrete white noise process:

$$\begin{aligned} E[w_i] &= 0 \\ E[w_i w_j^T] &= Q_{d_i}\delta_{ij}, \end{aligned} \tag{76}$$

and where $K = \exp(-\Delta t/\tau) < 1$. This is known as a 1st-order Markov process (MP) or an exponentially correlated random variable (ECRV). The mean or expected value of this process is given by

$$\boxed{\bar{x}_{i+1} = K\bar{x}_i, \quad \bar{x}_0 = 0} \tag{77}$$

and the variance is given by

$$\boxed{P_{i+1} = K^2 P_i + Q_{d_i}} \tag{78}$$

If a steady-state variance P_{ss} is desired, the noise variance must be set equal to

$$Q_{d_i} = P_{ss}(1 - K^2) = P_{ss}(1 - \exp[-2\Delta t/\tau]) \tag{79}$$

Also note that

$$E[x_{i+1}x_i] = E[(Kx_i + w_i)x_i] = KE[x_i^2] = KP_i = \exp(-\Delta t/\tau)P_i \tag{80}$$

The correlation between x_{i+1} and x_i is an exponential function of the time increment Δt between them.

1.7.6 A Scalar Continuous 1st-Order Stochastic Process

Suppose we have a scalar continuous linear system described by the equation

$$\dot{x} = -x/\tau + w \tag{81}$$

where w is a continuous white noise process

$$\begin{aligned} E[w(t)] &= 0 \\ E[w(t)w(t')^T] &= Q\delta(t-t'), \end{aligned} \tag{82}$$

and where τ is a known time-constant. This is known as a 1st-order Markov process (MP) or an exponentially correlated random variable (ECRV). The mean or expected value of this process is given by

$$\boxed{\dot{\bar{x}} = -\bar{x}/\tau = 0, \quad \bar{x}_0 = given} \tag{83}$$

and the variance is given by

$$\boxed{\dot{P} = -2P(t)/\tau + Q, \quad P_0 = given}$$

If a steady-state variance P_{ss} is desired, the noise variance must be set equal to

$$Q = 2P_{ss}/\tau \tag{84}$$

1.8 *Derivation* of the Time-Propagation Equations for the Mean and Covariance of a Random Process

1.8.1 Discrete Non-Linear Systems

A discrete-time nonlinear stochastic process, can be described by the following difference equation,

$$\boldsymbol{x}_{i+1} = \boldsymbol{h}(\boldsymbol{x}_i, \boldsymbol{w}_{d,i}), \quad \boldsymbol{x}_0 \sim N(\bar{\boldsymbol{x}}_0,\, P_0) \tag{85}$$

where $\boldsymbol{x}_i$ is an n-dimensional state vector, and where $\boldsymbol{w}_{d,i}$ is an m-dimensional vector with mean and covariance defined as

$$\begin{aligned} E[\boldsymbol{w}_{d,i}] &= 0 \\ E[\boldsymbol{w}_{d,i}\boldsymbol{w}_{d,j}^T] &= Q_{d_i}\delta_{ij}, \end{aligned} \tag{86}$$

Under the appropriate conditions, this non-linear difference equation can be linearized about the mean values of $\boldsymbol{x}_i$ and $\boldsymbol{w}_{d,i}$

$$\boldsymbol{x}_{i+1} \approx \boldsymbol{h}(\bar{\boldsymbol{x}}_i, \bar{\boldsymbol{w}}_{d,i}) + H_i(\boldsymbol{x}_i - \bar{\boldsymbol{x}}) + \Gamma_i(\boldsymbol{w}_{d,i} - \bar{\boldsymbol{w}}_{d,i}) \tag{87}$$

where

$$H_i = \left.\frac{\partial \boldsymbol{h}_i(\boldsymbol{x}_i, \boldsymbol{w}_{d,i})}{\partial \boldsymbol{x}_i}\right|_{\bar{\boldsymbol{x}}, \bar{\boldsymbol{w}}_d}, \quad \Gamma_i = \left.\frac{\partial \boldsymbol{h}_i(\boldsymbol{x}_i, \boldsymbol{w}_{d,i})}{\partial \boldsymbol{w}_{d,i}}\right|_{\bar{\boldsymbol{x}}, \bar{\boldsymbol{w}}_d} \tag{88}$$

Taking the expectation of Eq. 87 produces the mean of $\boldsymbol{x}_{i+1}$

$$\boxed{\bar{\boldsymbol{x}}_{i+1} = \boldsymbol{h}(\bar{\boldsymbol{x}}_i, \boldsymbol{w}_{d,i}), \quad \bar{\boldsymbol{x}}_0 = given} \tag{89}$$

The covariance of $\boldsymbol{x}_{i+1}$is given by:

$$P_{i+1} = E[(\boldsymbol{x}_{i+1} - \bar{\boldsymbol{x}}_{i+1})(\boldsymbol{x}_{i+1} - \bar{\boldsymbol{x}}_{i+1})^T] \tag{90}$$

where $\boldsymbol{x}_{i+1}$ and $\bar{\boldsymbol{x}}_{i+1}$ are given by Eqs. 87 and 89. Substituting these into Eq. 90, multiplying out all the terms, and taking the expectation produces

$$\boxed{P_{i+1} = H_i P_i H_i^T + \Gamma_i Q_{d_i} \Gamma_i^T, \quad P_0 = given} \tag{91}$$

1.8.2 Continuous Linear and Non-Linear Systems

The derivation of the mean and covariance for continuous linear systems can be found in Maybeck, Chapter 4. For nonlinear systems, the mean and covariance propagation equations can be determined by first linearizing the system, and then using the results from the continuous linear system theory.

Alternately, if the continuous linear or continuous non-linear system is first discretized, the mean and covariance propagation equations can be determined using the results from the discrete linear and discrete nonlinear system theory.

1.9 Final Remarks

1.9.1 Numerical Integration of Continuous White Noise

Continuous white noise is a convenient artificial mathematical entity invented to provide a mathematical description of real-world *diffusion* processes. Consider one particular diffusion process, a random walk.

$$\dot{x} = w \tag{92}$$

where w is a *continuous* white noise process:

$$\begin{aligned} E[w(t)] &= 0 \\ E[w(t)w(t')^T] &= Q\delta(t-t'), \end{aligned} \tag{93}$$

Notice that while the variance of w at a given time is infinite (try implementing a random variable with infinite variance in a compute code!), the variance of x can be determined using simple linear stochastic system theory

$$P(t+\Delta t) = P_0 + Q\Delta t \tag{94}$$

In this case, Q is the diffusion coefficient and is often called the strength of the white noise or the power spectral density.

So if we want numerically integrate equation 92, how do we model the white noise? Since it is impossible to precisely model continuous white noise on a computer, we are free to model it in any way we choose provided that we end up with the same statistics, $P(t+\Delta t) = P_0 + Q\Delta t$.

If we replace the continuous white noise with a discrete white noise process $w_{d,i}$ such that $w_{d,i}$is a constant for $t_i < t < t_{i+1}$, we have

$$\dot{x} = w_{d,i} \tag{95}$$

where

$$\begin{aligned} E[w_{d,i}] &= 0 \\ E[w_{d,i}w_{d,j}^T] &= Q_{d_i}\delta_{ij}, \end{aligned} \tag{96}$$

Notice that in this model, $w_{d,i}$ has a finite variance. Using stochastic linear system theory, the variance of x is now given by

$$P_{i+1} = P_i + Q_{d_i}\Delta t^2 \tag{97}$$

For this model of the white noise to produce the same statistics as the continuous white noise we must have

$$Q_{d_i} = \frac{Q}{\Delta t} \tag{98}$$

If the stochastic process of interest is more complex, but still linear

$$\dot{\boldsymbol{x}} = A\boldsymbol{x} + B\boldsymbol{w} \tag{99}$$

we know from stochastic linear system theory that the variance propagates according to

$$\dot{P} = AP + PA^T + BQB^T \tag{100}$$

If we consider a small time-step, this can be approximated to first order as

$$P_{i+1} = P_i + AP_i\Delta t + PA_i^T\Delta t + BQB^T\Delta t \tag{101}$$

Now if we replace the continuous white noise with a discrete white noise process,

$$\dot{\boldsymbol{x}} = A\boldsymbol{x} + B\boldsymbol{w}_{d,i} \tag{102}$$

the covariance of $\boldsymbol{x}$ is given

$$P_{i+1} = e^{A\Delta t}P_i e^{A^T\Delta t} + B_d Q_{d,i} B_d^T \tag{103}$$

where

$$B_d = \int_{t_i}^{t_{i+1}} e^{A(t_{i+1}-\tau)}Bd\tau \tag{104}$$

If we again consider a small time-step, this can be approximated to first order as

$$P_{i+1} = P_i + AP\Delta t + PA^T\Delta t + BQ_{d,i}B^T\Delta t^2 \tag{105}$$

Thus, to produce the same statistics, we again need to set $Q_{d,i} = Q/\Delta t$.

The same analysis will hold for a non-linear system $\dot{\boldsymbol{x}} = \boldsymbol{f}(\boldsymbol{x}, \boldsymbol{w})$, i.e. if we replace continuous white noise of strength Q with a discrete white noise process with covariance $Q_{d,i} = Q/\Delta t$, the resulting statistics will be the same to first order.

Is it important that the results are accurate only to first order? In most applications the white noise is a very small component of the dynamics. So a first order approximation of the effect of the white noise is probably acceptable. If higher order accuracy is required, there are special Runge-Kutta integrators that can model the white noise to higher order (see course website), but the added complexity may not be justified.

Finally, and perhaps most importantly, in real-world applications, the uncertainties and errors in the dynamics model are probably not white noise to begin with! The errors may be very small random sinusoidal terms, or very small discrete changes in acceleration, or small slowly changing biases. It is the responsibility of the engineer is to evaluate the effect of these uncertainties in terms of how they contribute to the growth of the state covariance and then to model them as a simple *diffusion* process driven by continuous white noise with strength Q. In this case, implementing a first-order approximation of white noise in a computer code will be quite acceptable.

1.9.2 One More Subtle Point

There are two common ways to model continuous white noise as a discrete white noise process. Consider the continuous-time stochastic system are given by

$$\dot{\boldsymbol{x}} = \boldsymbol{f}(\boldsymbol{x}) + \boldsymbol{w} \tag{106}$$

$$\dot{P} = FP + PF^T + Q \tag{107}$$

where

$$F_i = \left.\frac{\partial \boldsymbol{f}(\boldsymbol{x},\boldsymbol{w})}{\partial \boldsymbol{x}}\right|_{\bar{\boldsymbol{x}}_i,\bar{\boldsymbol{w}}_i} \tag{108}$$

To first order we have

$$P_{i+1} = P_i + F_i P_i \Delta t + P F_i^T \Delta t + Q \Delta t \tag{109}$$

Notice that the continuous noise has units equal to the state per unit time, and the strength of the noise has units equal to the state squared per unit time.

The first approach to modeling the effects of continuous noise is to replace it with a discrete noise process (as we did in the previous section) where $\boldsymbol{w}_{d,i}$ has the same units of the continuous noise, state per unit time, and $Q_{d,i}$ has units of the state squared per time squared. To first order, we have

$$\dot{\boldsymbol{x}} = \boldsymbol{f}(\boldsymbol{x}) + \boldsymbol{w}_{d,i} \tag{110}$$

$$\boldsymbol{x}_{i+1} = \boldsymbol{x}_i + \boldsymbol{f}(\boldsymbol{x}_i)\Delta t + \boldsymbol{w}_{d,i}\Delta t \tag{111}$$

$$\bar{\boldsymbol{x}}_{i+1} = \bar{\boldsymbol{x}}_i + \boldsymbol{f}(\bar{\boldsymbol{x}}_i)\Delta t \tag{112}$$

$$\boldsymbol{x}_{i+1} - \bar{\boldsymbol{x}}_{i+1} = (I + F_i \Delta t)(\boldsymbol{x}_{i+1} - \bar{\boldsymbol{x}}_i) + \boldsymbol{w}_{d,i}\Delta t \tag{113}$$

$$P_{i+1} = (I + F_i \Delta t) P_i (I + F_i \Delta t)^T + Q_{d,i} \Delta t^2 \tag{114}$$

$$P_{i+1} \approx P_i + F_i P_i \Delta t + P_i F_i^T \Delta t + Q_{d,i} \Delta t^2 \tag{115}$$

For the covariance in Eq. 109 to match the covariance in Eq. 115, $Q_{d,i}$ must be set to $Q/\Delta t$.

The second approach to modeling the effects of continuous white noise is to first discretize the continuous system, and then add discrete white noise.

$$\boldsymbol{x}_{i+1} = \boldsymbol{x}_i + \boldsymbol{f}(\boldsymbol{x}_i)\Delta t + \boldsymbol{w}_{d2,i} \tag{116}$$

$$\bar{\boldsymbol{x}}_{i+1} = \bar{\boldsymbol{x}}_i + \boldsymbol{f}(\bar{\boldsymbol{x}}_i)\Delta t \tag{117}$$

$$\boldsymbol{x}_{i+1} - \bar{\boldsymbol{x}}_{i+1} = (I + F_i \Delta t)(\boldsymbol{x}_{i+1} - \bar{\boldsymbol{x}}_i) + \boldsymbol{w}_{d2,i} \tag{118}$$

$$P_{i+1} = (I + F_i \Delta t) P_i (I + F_i \Delta t)^T + Q_{d2,i} \tag{119}$$

$$P_{i+1} \approx P_i + F_i P_i \Delta t + P_i F_i^T \Delta t + Q_{d2,i} \tag{120}$$

Notice that in this case $\boldsymbol{w}_{d2,i}$ has units of the state and $Q_{d2,i}$ has units of the state squared. For the statistics to be the same as the continuous noise case, we must set the variance of $\boldsymbol{w}_{d2,i}$ to

$$Q_{d2,i} = Q \Delta t \tag{121}$$

The relationship between $Q_{d2,i}$and$Q_{d,i}$is

$$Q_{d2,i} = Q_{d,i} \Delta t^2 \tag{122}$$

The choice of which approach to use is dependent on the problem being solved.