Stochastic Adaptive Control (02421)

Lecture 2

Tobias K. S. Ritschel

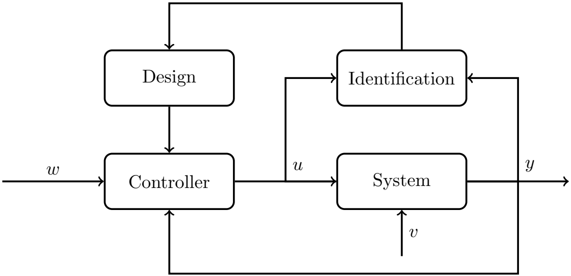
Section for Dynamical Systems

Department of Applied Mathematics and Computer Science

Technical University of Denmark

Lecture Plan

1. System theory
2. **Stochastics**
3. State estimation 1
4. State estimation 2
5. Optimal control 1
6. System identification 1 + adaptive control 1
7. External models + prediction
8. Optimal control 2
9. Optimal control 3
10. System identification 2
11. System identification 3 + model validation
12. System identification 4 + adaptive control 2
13. Adaptive control 3



Today's Agenda

- Follow-up from last lecture
- Probability theory
- Stochastic dynamical systems
- Discretization

Follow-up from Last Time: Exercise 4

Find eigenvalues and path to/from origin

$$x_{k+1} = \begin{bmatrix} 0 & 1 & 2 \\ 0 & 0 & 3 \\ 0 & 0 & 0 \end{bmatrix} x_k + \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} u_k = Ax_k + Bu_k \tag{1}$$

1) The eigenvalues indicate asymptotic stability:

$$\text{eig}(A) = \{0, 0, 0\} \tag{2}$$

2-3)

$$A^2 = \begin{bmatrix} 0 & 0 & 3 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \quad A^3 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \quad AB = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \tag{3}$$

Consequently, $U = \begin{bmatrix} 0 & 0 & 0 \end{bmatrix}$ drives the states to the origin, whereas $U = \begin{bmatrix} -3 & 0 \end{bmatrix}$ does so in the minimum number of steps.

4) No, the system cannot be driven from the origin to $\begin{bmatrix} 1 & 1 & 1 \end{bmatrix}^T$ because we cannot affect state 3.

Demonstration

Probability theory

Stochastics

Real systems are usually stochastic in nature

$$x_{k+1} = Ax_k + Bu_k + v_k, \qquad v_k \sim N(0, R_1), \tag{4}$$

$$y_k = Cx_k + Du_k + e_k, \qquad e_k \sim N(0, R_2) \tag{5}$$

Stochastic: being uncertain, described by a random distribution and cannot be predicted precisely.

Sources: measurements, model inaccuracy (unmodeled phenomena), unknown disturbances, unknown/time-varying parameters, etc.

Stochastic scalar variables

Stochastic variable

$$X \sim \mathcal{F}(\mathbf{p}) \tag{6}$$

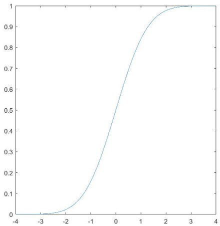
Cumulative distribution function $F_X(y)$ (cdf)

$$F_X(y) = \Pr\{X \leq y\} \in [0,1], \quad \Pr\{a \leq X \leq b\} = F_X(b) - F_X(a) \tag{7}$$

Probability density function $f_X(z) \geq 0$ (pdf)

$$F_X(y) = \int_{-\infty}^{y} f_X(z)\,\mathrm{d}z, \tag{8}$$

$$F(-\infty) = 0, \quad F(\infty) = 1 \tag{9}$$



Confidence Interval

$1 - p$ confidence interval $CI(p)$

$$\Pr\{a \leq X \leq b\} = 1 - p \tag{10}$$

Confidence interval based on inverse cdf

$$\Pr\{X \leq a\} = p/2 \quad \text{or} \quad \Pr\{X \leq b\} = 1 - p/2 \tag{11}$$

$$CI(p) = [F_X^{-1}(p/2), F_X^{-1}(1 - p/2)] \tag{12}$$

Use Matlab routines (or look-up tables) to compute $F^{-1}(p/2)$

$$X \in m_X \pm \sigma_X F^{-1}(p/2) \tag{13}$$

Example: Let $X \sim N(10, 4)$. Then, a 95% CI is

$$10 - 2 \cdot 1.96 \leq X \leq 10 + 2 \cdot 1.96 \text{ or } 6.08 \leq X \leq 13.92 \tag{14}$$

Moments

For a real function $g(X)$

$$\textbf{Nth moment of g(X):}\ \mathbb{E}[g^n(X)] = \int_\Omega g^n(x) f(x)\,\mathrm{d}x \tag{15}$$

Moments represent certain properties of stochastic variables.

$$\textbf{Mean (1st moment):}\ \mathbb{E}[X] = m_x \tag{16}$$

$$\begin{aligned}\textbf{Variance (2nd central moment):}\ \mathrm{Var}(X) &= \mathbb{E}[(X - m_x)^2] \\ &= \mathbb{E}[X^2] - \mathbb{E}[X]^2 = \sigma_x^2 \end{aligned} \tag{17}$$

$$\textbf{Skewness (std.* 3rd central moment):}\ \mathbb{E}[(X - m_x)^3]/\sigma_x^3 \tag{18}$$

*Standardized.

Moments - Sample Moments

Let $\{x_i\}_{i=1}^N$ be samples of X

Estimates of first and second order moments

$$\mathbb{E}[X] = \sum_{i=1}^{N} \frac{x_i}{N} \tag{19}$$

$$\mathrm{Var}(X) = \sum_{i=1}^{N} \frac{(x_i - \mathbb{E}[X])^2}{N} \tag{20}$$

Unbiased estimate of variance

$$\mathrm{Var}(X) = \sum_{i=1}^{N} \frac{(x_i - \mathbb{E}[X])^2}{N-1} \tag{21}$$

$$\tag{22}$$

Probabilities: Joint probability and independence

Marginal probability that a single statement $(X \leq x)$ is true

$$\Pr\{X \leq x\} = F_X(x) \tag{23}$$

Joint probability that two (or more) statements are true

$$\Pr\{X \leq x, Y \leq y\} = F_{X,Y}(x,y) \tag{24}$$

Compute the **marginal distribution** from the joint distribution

$$f_X(x) = \int_{\Omega_y} f_{X,Y}(x,y)\,\mathrm{d}y \tag{25}$$

Joint distributions for **independent** variables

$$F_{X,Y}(x,y) = F_X(x)F_Y(y), \quad f_{X,Y}(x,y) = f_X(x)f_Y(y) \tag{26}$$

Covariance

Covariance is a measure of how two stochastic variables varies relatively to each other

$$\mathrm{Cov}(X,Y) = \mathbb{E}[(X - m_x)(Y - m_y)] \tag{27}$$

Variance is covariance between the same variable

$$\mathrm{Var}(X) = \mathrm{Cov}(X,X) \tag{28}$$

Correlation coefficient

$$\rho = \frac{\mathrm{Cov}(X,Y)}{\sqrt{\mathrm{Var}(X)\,\mathrm{Var}(Y)}}, \quad -1 \leq \rho \leq 1 \tag{29}$$

Covariance of independent variables

$$\mathrm{Cov}(X,Y) = \rho = 0 \tag{30}$$

Note: The reverse if not true

Conditional distribution – Bayes' theorem

Conditional probability and Bayes' theorem

$$\Pr(A|B)\Pr(B) = \Pr(A,B) = \Pr(B|A)\Pr(A), \tag{31}$$

$$\Pr\{X \le x|Y \le y\}\Pr\{Y \le y\} = \Pr\{X \le x, Y \le y\} \tag{32}$$

$$= \Pr\{Y \le y|X \le x\}\Pr\{X \le x\} \tag{33}$$

Conditional probability density function

$$f_{X|Y}(x|y)f_Y(y) = f_{X,Y}(x,y) = f_{Y|X}(y|x)f_X(x) \tag{34}$$

The same can be done for the moments if $\mathrm{Var}(X|Y) < \infty$ exists

$$\mathbb{E}[X|Y] = m_{x|y} = \int_{\Omega_x} x f_{X|Y}(x|y)\,\mathrm{d}x \tag{35}$$

$$\mathrm{Var}(X|Y) = \mathbb{E}[(X - m_{x|y})^2|Y] \tag{36}$$

Stochastic Vectors

Vector-valued random variables

$$\mathbf{X} = \left[X_1, \ldots, X_n\right]^T \tag{37}$$

$$\text{cdf: } F_{\mathbf{X}}(\mathbf{x}) = \Pr(X_1 \leq x_1, \ldots, X_n \leq x_n,) \tag{38}$$

$$\text{marginal cdf: } F_{X_1}(x_1) = \Pr(X_1 \leq x_1) \tag{39}$$

1st and 2nd order moments

$$\mathbf{m}_x = \mathbb{E}[\mathbf{X}] = \left[\mathbb{E}[X_1], \ldots, \mathbb{E}[X_n]\right]^T \tag{40}$$

$$P_x = P_x^T = \mathrm{Cov}(\mathbf{X}) = \mathbb{E}[(\mathbf{X} - \mathbf{m}_x)(\mathbf{X} - \mathbf{m}_x)^T] \succeq 0 \tag{41}$$

Positive semi-definiteness ($\succeq$) means that $x^T P_x x \geq 0$.

Covariance matrices are diagonalizable, e.g., for $n = 2$

$$P_x = \begin{bmatrix} \mathrm{Var}(X_1) & \mathrm{Cov}(X_1, X_2) \\ \mathrm{Cov}(X_2, X_1) & \mathrm{Var}(X_2) \end{bmatrix} \tag{42}$$

Mathematical properties of Moments: 1st and 2nd

Let A and m be a constant matrix and vector

Expectations

$$\begin{aligned}
\mathbb{E}[\mathbf{X}+m] &= \mathbb{E}[\mathbf{X}]+m & (43)\\
\mathbb{E}[A\mathbf{X}] &= A\mathbb{E}[\mathbf{X}] & (44)\\
\mathbb{E}[\mathbf{X}+\mathbf{Y}] &= \mathbb{E}[\mathbf{X}]+\mathbb{E}[\mathbf{Y}] & (45)\\
\mathbb{E}[\mathbf{X}^T A\mathbf{X}] &= \mathrm{Tr}(A\,\mathrm{Cov}(\mathbf{X}))+\mathbb{E}[\mathbf{X}]^T A\mathbb{E}[\mathbf{X}] & (46)
\end{aligned}$$

Covariances

$$\begin{aligned}
\mathrm{Cov}(\mathbf{X}) &= \mathbb{E}[\mathbf{X}\mathbf{X}^T]-\mathbb{E}[\mathbf{X}]\mathbb{E}[\mathbf{X}]^T & (47)\\
\mathrm{Cov}(\mathbf{X}+m) &= \mathrm{Cov}(\mathbf{X}) & (48)\\
\mathrm{Cov}(A\mathbf{X}) &= A\,\mathrm{Cov}(\mathbf{X})A^T & (49)\\
\mathrm{Cov}(\mathbf{X}+\mathbf{Y}) &= \mathrm{Cov}(\mathbf{X})+\mathrm{Cov}(\mathbf{Y})+\mathrm{Cov}(\mathbf{X},\mathbf{Y})+\mathrm{Cov}(\mathbf{X},\mathbf{Y})^T & (50)
\end{aligned}$$

Vector Covariance and Variance

Further covariances

$$\mathrm{Cov}(\mathbf{X},\mathbf{Y}) = \mathbb{E}[(\mathbf{X}-m_x)(\mathbf{Y}-m_y)^T] \tag{51}$$

$$\mathrm{Cov}(\mathbf{X},\mathbf{X}) = \mathrm{Cov}(\mathbf{X}) = P_x \tag{52}$$

$$\mathrm{Cov}(\mathbf{Y},\mathbf{X}) = \mathrm{Cov}(\mathbf{X},\mathbf{Y})^T \tag{53}$$

$$\mathrm{Cov}(A\mathbf{X},\mathbf{Y}) = A\,\mathrm{Cov}(\mathbf{X},\mathbf{Y}) \tag{54}$$

$$\mathrm{Cov}(\mathbf{X},A\mathbf{Y}) = \mathrm{Cov}(\mathbf{X},\mathbf{Y})A^T \tag{55}$$

$$\mathrm{Cov}(\mathbf{X}+\mathbf{V},\mathbf{Y}) = \mathrm{Cov}(\mathbf{X},\mathbf{Y}) + \mathrm{Cov}(\mathbf{V},\mathbf{Y}) \tag{56}$$

Principal directions of the variance (PCA)

$$[\Lambda,\mathbf{V}] = \mathrm{eig}(P_x) \tag{57}$$

$$P_x\mathbf{V}_i = \lambda_i\mathbf{V}_i \tag{58}$$

The columns in **V** indicate the main directions of the variation and the elements of Λ indicate the associated variance

Different Distributions - Gaussian and χ^2

Gaussian/normal distribution

$$X \sim N(m_x, \sigma_x^2) \tag{59}$$

$$Y = \frac{X - m_x}{\sigma_x} \sim N(0,1) \quad \text{standard Gaussian} \tag{60}$$

$$f_X(x) = \frac{1}{\sqrt{2\pi}\sigma_x} \exp\left(-\frac{(x-m_x)^2}{2\sigma_x^2}\right) \tag{61}$$

$$F_X(x) = F_Y\left(\frac{x-m_x}{\sigma_x}\right) \tag{62}$$

χ^2-distribution

$$X = \sum_{i=1}^{n} \psi_i^2 \sim \chi^2(n), \quad \psi_i \sim N(0,1), \quad \psi_i \perp \psi_j \tag{63}$$

$$f(x) = \frac{1}{\Gamma(n/2)} x^{n/2-1} \exp\left(-\frac{x}{2}\right) \tag{64}$$

$$\mathbb{E}[X] = n \quad \mathrm{Var}(X) = 2n \tag{65}$$

Different Distributions - Gamma

Gamma distribution ($\chi^2(n) = \Gamma(n/2, 2)$)

$$X \sim \Gamma(k, \theta), \quad 0 < X < \infty \tag{66}$$

$$f_X(x) = \frac{1}{\Gamma(k)\theta^k} x^{k-1} \exp\left(-\frac{x}{\theta}\right) \tag{67}$$

$$\mathbb{E}[X] = k\theta, \quad \mathrm{Var}(X) = k\theta^2 \tag{68}$$

Gamma function

$$\Gamma(k) = \int_0^\infty t^{k-1} e^{-t} dt \tag{69}$$

$$\Gamma(k+1) = k\Gamma(k) \tag{70}$$

$$\Gamma(1) = 1, \quad \Gamma\left(\frac{1}{2}\right) = \sqrt{\pi} \tag{71}$$

Erlang distribution (Gamma distribution for integer values of k)

$$\Gamma(k) = (k-1)!, \quad \Gamma\left(k + \frac{1}{2}\right) = \frac{(2k-1)!}{2^k}\sqrt{\pi} \tag{72}$$

Other Related Distributions

The F-distribution

$$X = \frac{Zm}{Yn} \sim F(n,m) \tag{73}$$

$$Z \sim \chi^2(n), \quad Y \sim \chi^2(m), \quad Z \perp Y \tag{74}$$

Student's t-distribution

$$X = \frac{Z}{\sqrt{Y}}\sqrt{n} \sim t(n) \tag{75}$$

$$Z \in N(0,1), \quad Y \sim \chi^2(n), \quad Z \perp Y \tag{76}$$

The Rayleigh distribution:

$$X = \sqrt{Y_1^2 + Y_2^2} \sim Ray(\sigma_y^2), \quad Y_i \sim N_{iid}(0, \sigma_y^2) \tag{77}$$

Generate random variables in Matlab

Stochastic variable change (use Matlab's `randn`)

$$X \sim N(m, P), \quad Z \sim N(0, I) \tag{78}$$
$$X = SZ + m \tag{79}$$
$$P = SS^T \tag{80}$$

Cholesky factorization

$$S^T = \mathrm{chol}(P) \tag{81}$$

Eigenvalue decomposition

$$Pv_i = \lambda_i v_i \rightarrow PV = VD \tag{82}$$
$$S = V\sqrt{D} \tag{83}$$

Alternative: Use Matlab's `mvnrnd`

Hint: Don't use routines with limited functionality, e.g., `normrnd`

Stochastics in Matlab

Sampling from distributions in Matlab

Stochastic discrete-time systems

A Stochastic System

We will now extend the discrete-time (deterministic) systems

$$x_{k+1} = Ax_k + Bu_k, \tag{84a}$$
$$y_k = Cx_k + Du_k \tag{84b}$$

to discrete-time stochastic systems of the form

$$x_{k+1} = Ax_k + Bu_k + Gv_k, \tag{85a}$$
$$y_k = Cx_k + Du_k + Fe_k \tag{85b}$$

Two sources of uncertainty/noise

- $\{v_k, k \in \mathbb{N}\}$ denotes the process noise
- $\{e_k, k \in \mathbb{N}\}$ denotes the measurement/sampling noise

x_k is a random variable and the evolution of the system is a stochastic process

Stochastic Processes

A stochastic process can be described using a marginal cdf or pdf

$$F_{X_t}(x_t, t) = \Pr\{X_t \leq x_t\} \tag{86}$$

$$f_{X_t}(x_t, t) = \nabla_{x_t} F_{X_t}(x_t, t) \tag{87}$$

or if the different times are related, using joint probabilities

$$F_{X_t, X_s}(x_t, x_s, t, s) = \Pr\{X_t \leq x_t, X_s \leq x_s\} \tag{88}$$

Properties of Stochastic processes

Mean

$$m_x(t) = \mathbb{E}[x(t)] = \int_{-\infty}^{\infty} z f_{x(t)}(z)\, \mathrm{d}z, \tag{89}$$

Variance

$$P_x(t) = \mathrm{Var}(x(t)) = \mathbb{E}\Big[\big(x(t) - \mathbb{E}[x(t)]\big)\big(x(t) - \mathbb{E}[x(t)]\big)^T\Big], \tag{90}$$

Auto-covariance

$$r_x(t_1, t_2) = \mathrm{Cov}(x(t_1), x(t_2)) = \mathbb{E}\Big[\big(x(t_1) - \mathbb{E}[x(t_1)]\big)\big(x(t_2) - \mathbb{E}[x(t_2)]\big)^T\Big], \tag{91}$$

Note that $r_x(t,t) = P_x(t)$

Auto-covariance and auto-correlation

Auto-correlation function

$$\rho_x(t_1, t_2) = \frac{r_x(t_1, t_2)}{\sqrt{P_x(t_1)P_x(t_2)}} \tag{92}$$

For stationary processes, only the time difference $\tau = t_1 - t_2$ is relevant

$$r_x(\tau) = \mathrm{Cov}\left[x(t), x(t+\tau)\right] \tag{93a}$$

$$\rho_x(\tau) = \frac{r_x(\tau)}{P_x(\tau)} \tag{93b}$$

Stationarity

Steady state of deterministic systems: The state does not change in time

Stationary distributions of stochastic systems: The state distribution does not change in time

Strong stationarity

$$f_{x(t_1),\ldots,x(t_n)}(z_1,\ldots,z_n) = f_{x(t_1+\Delta t),\ldots,x(t_n+\Delta t)}(z_1,\ldots,z_n), \tag{94}$$

for any $n \in \mathbb{N}$ and $\Delta t \in \mathbb{R}$.

Weak stationarity: The first two moments (mean and covariance) do not change in time

Normal process and Markov process

Normal process: Any probability density function $f_{x(t_1),\dots,x(t_n)}(z_1,\dots,z_n)$ is a multivariate normal distribution for any $n \in \mathbb{N}$

Probability density function with mean μ and covariance Σ

$$f_Y(y) = \frac{1}{(2\pi)^{n/2}\sqrt{\det(\Sigma)}} \exp\left(-\frac{1}{2}(y-\mu)^T\Sigma^{-1}(y-\mu)\right) \tag{95}$$

Markov process: For any $t_1 < t_2 < \cdots < t_n$, the distribution of $x(t_n)$ given $(x(t_1),\dots,x(t_{n-1}))$ is the same as the distribution of $x(t_n)$ given $x(t_{n-1})$

$$\Pr\Big(x(t_n) \le x \mid x(t_{n-1}),\dots,x(t_1)\Big) = \Pr\Big(x(t_n) \le x \mid x(t_{n-1})\Big) \tag{96}$$

Stochastic State-Space Models

Discrete-time system

$$x_{k+1} = Ax_k + Bu_k + Gv_k, \quad v_k \sim N(\mu_v, R_1) \tag{97a}$$

$$y_k = Cx_k + Du_k + Fe_k, \quad e_k \sim N(\mu_e, R_2) \tag{97b}$$

Mean and covariance

$$\boxed{\begin{aligned} \mu_{k+1} &= A\mu_k + Bu_k + G\mu_v, \quad \mu_0 = \mathbb{E}[x_0], && (98a) \\ P_{k+1} &= AP_kA^T + GR_1G^T, \quad P_0 = \mathrm{Cov}(x_0) && (98b) \end{aligned}}$$

Note that u_k is deterministic.

Stationary mean and covariance

$$\boxed{\begin{aligned} \mu_\infty &= A\mu_\infty + Bu_\infty + G\mu_v, && (99a) \\ P_\infty &= AP_\infty A^T + GR_1G^T && (99b) \end{aligned}}$$

Stationary auto-covariance (if A has full-rank and the eigenvalues lie within the unit circle)

$$r_{x,\infty}(\tau) = A^\tau P_\infty \tag{100}$$

Continuous-Time Stochastic Processes (SDE)

Continuous-time system

$$\dot{x}(t) = f(t, x(t), u(t)) \tag{101}$$

First attempt at stochastic differential equation

$$\dot{x}(t) = f(t, x(t), u(t)) + g(t, x(t), u(t))v(t) \tag{102}$$

Process noise

- $v(t) \perp v(s)$ for any $t \neq s$ (independence)
- $v(t)$ is continuous and has bounded variance
- $\mathbb{E}[v(t)] = 0$ (zero-mean)

Theorem 4.1 in Chapter 3 of the book "Stochastic Control Theory" by Åström (1970): $\mathbb{E}[v^2(t)] = 0$

Continuous-Time Stochastic Processes (SDE)

Stochastic difference equation

$$x(t+\Delta t) - x(t) = f(t, x(t))\Delta t + g(t, x(t))v(t)\Delta t + o(\Delta t). \tag{103}$$

Replace $v(t)\Delta t$ with $\Delta w(t) = w(t+\Delta t) - w(t)$, which has stationary independent zero-mean increments (Wiener process)

$$\Delta x(t) = f(t, x(t))\Delta t + g(t, x(t))\Delta w(t) + o(\Delta t). \tag{104}$$

Take the limit $\Delta t \to 0$

$$\mathrm{d}x(t) = f(t, x(t))\,\mathrm{d}t + g(t, x(t))\,\mathrm{d}w(t) \tag{105}$$

$$x(t) = x(t_0) + \int_{t_0}^{t} f(\tau, x(\tau))\,\mathrm{d}\tau + \int_{t_0}^{t} g(\tau, x(\tau))\,\mathrm{d}w(\tau) \tag{106}$$

Continuous-Time Stochastic Processes (SDE)

The two first conditional moments of the difference process

$$\mathbb{E}\Big[\Delta x(t) \mid x(t)\Big] = f(t, x(t))\Delta t + o(\Delta t) \tag{107a}$$

$$\mathrm{Var}\Big(\Delta x(t) \mid x(t)\Big) = g^2(t, x(t))\Delta t + o(\Delta t) \tag{107b}$$

Variance of process noise increment

$$\mathbb{E}\Big[\Delta w^2(t)\Big] = \Delta t \tag{108}$$

Note that the variance is proportional to Δt and not Δt^2

Distribution of process noise increment (increment of Wiener process)

$$\Delta w(t) = w(t + \Delta t) - w(t) \sim N(0, \Delta t) \tag{109}$$

Linear Stochastic Differential Equations

Linear stochastic differential equations

$$\mathrm{d}x(t) = (A(t)x(t) + B(t)u(t))\,\mathrm{d}t + G(t)\,\mathrm{d}w(t), \quad x(t_0) \sim N(m_0, P_0) \tag{110}$$

$A(t)$ and $B(t)$ are continuous functions of time

State expectation

$$\mathbb{E}[x(t)] = \mathbb{E}[x_0] + \mathbb{E}\left[\int_{t_0}^{t} A(\tau)x(\tau) + B(\tau)u(\tau)\,\mathrm{d}\tau\right] + \mathbb{E}\left[\int_{t_0}^{t} G(\tau)\,\mathrm{d}w(\tau)\right] \tag{111}$$

$$= \mathbb{E}[x_0] + \int_{t_0}^{t} A(\tau)\mathbb{E}[x(\tau)] + B(\tau)u(\tau)\,\mathrm{d}\tau = m_x(t) \tag{112}$$

Expected value

$$\dot{m}_x(t) = Am_x(t) + B(t)u(t), \quad m_x(t_0) = m_0. \tag{113}$$

Linear Stochastic Differential Equations

State-transition matrix

$$\frac{\partial \Phi(t;t_0)}{\partial t} = A(t)\Phi(t;t_0), \qquad \Phi(t_0;t_0) = I. \tag{114}$$

Auto-covariance of x $(s \geq t)$

$$R(s,t) = \mathrm{Cov}(x(s), x(t)) = \Phi(s,t)P(t) \tag{115}$$

Covariance

$$\dot{P}(t) = A(t)P(t) + P(t)A^T(t) + G(t)G^T(t), \qquad P(t_0) = P_0, \tag{116}$$

Discretization

Linear stochastic differential equation

Linear continuous-time state space model

$$dx(t) = (Ax(t) + Bu(t))\,dt + G\,dw(t), \qquad dw(t) \sim N(0, I\,dt), \tag{117a}$$
$$y(t) = Cx(t) + Du(t) + Fe(t), \qquad e(t) \sim N(m_e, R_e) \tag{117b}$$

Zero-order-hold parametrization of manipulated inputs

$$u(t) = u_k, \qquad t \in [t_k, t_{k+1}[\tag{118}$$

Approximation of process noise (not rigorous)

$$dw(t) = \tilde{w}(t)\,dt, \qquad \tilde{w}(t) \sim N(0, I) \tag{119}$$

Analytical solution

$$x(t_{k+1}) = e^{A(t_{k+1}-t_k)}x(t_k) + \int_{t_k}^{t_{k+1}} e^{A(t_{k+1}-\tau)}Bu(t_k)\,d\tau + v(t_k), \tag{120a}$$
$$y(t_k) = Cx(t_k) + Du(t_k) + Fe(t_k), \quad e(t_k) \sim N(m_e, R_e) \tag{120b}$$

Zero-order-hold parametrization and approximation

Discrete-time process noise

$$v(t_k) = \int_{t_k}^{t_{k+1}} e^{A(t_{k+1}-\tau)} G \tilde{w}(\tau) \, \mathrm{d}\tau \tag{121a}$$

Mean

$$\mathbb{E}[v(t_k)] = \int_{t_k}^{t_{k+1}} e^{A(t_{k+1}-\tau)} G \mathbb{E}[\tilde{w}(\tau)] \, \mathrm{d}\tau = 0 \tag{122}$$

Zero-order-hold parametrization and approximation

Covariance

$$\mathrm{Cov}(v(t_k)) = \mathbb{E}[v(t_k)v^T(t_k)] \tag{123}$$

$$= \mathbb{E}\left[\int_{t_k}^{t_{k+1}} e^{A(t_{k+1}-\tau)} G \tilde{w}(\tau)\, \mathrm{d}\tau \int_{t_k}^{t_{k+1}} \tilde{w}^T(s) G^T e^{A^T(t_{k+1}-s)}\, \mathrm{d}s \right] \tag{124}$$

$$= \int_{t_k}^{t_{k+1}} \int_{t_k}^{t_{k+1}} e^{A(t_{k+1}-\tau)} G \mathbb{E}\left[\tilde{w}(\tau)\tilde{w}^T(s) \right] G^T e^{A^T(t_{k+1}-s)}\, \mathrm{d}\tau\, \mathrm{d}s \tag{125}$$

$$= \int_{t_k}^{t_{k+1}} e^{A(t_{k+1}-\tau)} G \mathbb{E}\left[\tilde{w}(\tau)\tilde{w}^T(\tau) \right] G^T e^{A^T(t_{k+1}-\tau)}\, \mathrm{d}\tau \tag{126}$$

$$= \int_{t_k}^{t_{k+1}} e^{A(t_{k+1}-\tau)} G G^T e^{A^T(t_{k+1}-\tau)}\, \mathrm{d}\tau \tag{127}$$

Stochastic discrete-time state space models

Linear discrete-time stochastic state space model

$$x_{k+1} = A_d x_k + B_d u_k + v_k, \qquad v_k \sim N(0, R_1), \tag{128}$$

$$y_k = C_d x_k + D_d u_k + e_k, \qquad e_k \sim N(0, R_2) \tag{129}$$

System matrices in the state equation

$$\begin{bmatrix} A_d & B_d \\ 0 & I \end{bmatrix} = \exp\left(\begin{bmatrix} A & B \\ 0 & 0 \end{bmatrix} T_s \right), \tag{130}$$

$$\begin{bmatrix} A_d & \tilde{R}_1 \\ 0 & A_d^{-T} \end{bmatrix} = \exp\left(\begin{bmatrix} A & GG^T \\ 0 & -A^T \end{bmatrix} T_s \right), \qquad R_1 = \tilde{R}_1 A_d^T \tag{131}$$

$$\begin{bmatrix} A_d^{-1} & \tilde{R}_1^T \\ 0 & A_d^T \end{bmatrix} = \exp\left(\begin{bmatrix} -A & GG^T \\ 0 & A^T \end{bmatrix} T_s \right), \qquad R_1 = A_d \tilde{R}_1^T \tag{132}$$

System matrices in the measurement equation

$$C_d = C, \qquad D_d = D, \qquad R_2 = F R_e F^T \tag{133}$$

Proof of discretization

Exponential form

$$M = \begin{bmatrix} X & Y \\ 0 & Z \end{bmatrix} = \exp\left(\begin{bmatrix} F & G \\ 0 & H \end{bmatrix} t\right) = \exp(Kt) \tag{134}$$

Differential form

$$\dot{M} = KM, \quad M(t_0) = I \tag{135}$$

Individual differential equations

$$\dot{X} = FX, \qquad X(t_0) = I, \tag{136}$$

$$\dot{Y} = FY + GZ, \qquad Y(t_0) = 0, \tag{137}$$

$$\dot{Z} = HZ, \qquad Z(t_0) = I \tag{138}$$

Solutions

$$X = e^{F(t-t_0)} X(t_0) = e^{F(t-t_0)}, \tag{139}$$

$$Z = e^{H(t-t_0)} Z(t_0) = e^{H(t-t_0)}, \tag{140}$$

$$Y = e^{F(t-t_0)} Y(t_0) + \int_{t_0}^{t} e^{F(t-\tau)} G e^{H(\tau-t_0)} Z(t_0)\, \mathrm{d}\tau = \int_{t_0}^{t} e^{F(t-\tau)} G e^{H(\tau-t_0)}\, \mathrm{d}\tau$$

Proof of discretization

Let $F = A$, $H = -A^T$, $G = GG^T$, $t_0 = t_k$, and $t = t_{k+1}$ $(t_{k+1} - t_k = T_s)$

$$X = e^{AT_s} = A_d, \tag{141}$$

$$Z = e^{-A^T T_s} = A_d^{-T}, \tag{142}$$

$$Y = \int_{t_k}^{t_{k+1}} e^{A(t_{k+1}-\tau)} GG^T e^{A^T(t_{k+1}-\tau)}\, \mathrm{d}\tau\, e^{-A^T T_s} = \tilde{R}_1 = R_1 A_d^T \tag{143}$$

The proof of the other approach is similar, but has one more step:
A change of variables in the integral

Questions

Questions?