CALIFORNIA INSTITUTE OF TECHNOLOGY
Control and Dynamical Systems

CDS 112/Ae 103a

R. M. Murray **Stochastic Systems** 7 February 2022

This set of lectures provides a brief introduction to stochastic systems.

Reading:

- Friedland, Chapter 10

1 Introduction

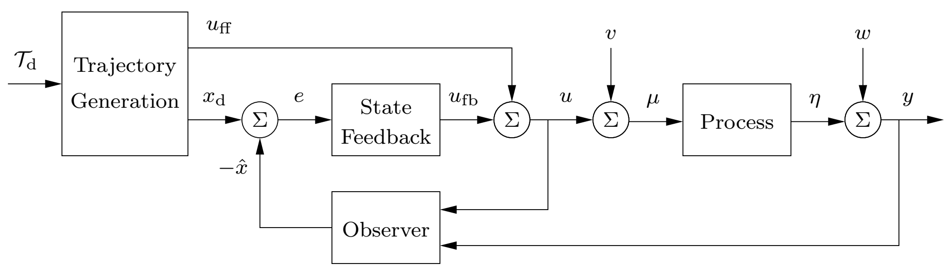


Figure 1: Block diagram of a basic feedback loop.

$$
\begin{aligned}
\dot{x} &= f(x,u), & x(0) &= x_0 \in \mathbb{R}^n, \\
y &= h(x), & u(t) &\in \mathbb{R}^m, \\
& & y &\in \mathbb{R}^p,
\end{aligned}
\quad \rightarrow \quad
\begin{aligned}
& & X(0) &\sim \mathcal{N}(x_0, \Sigma_X), \\
\dot{X} &= F(X,u,V), & u &\in \mathbb{R}^m, \\
Y &= h(X) + W & V(t) &\sim \mathcal{N}(0, \Sigma_V), \\
& & W(t) &\sim \mathcal{N}(0, \Sigma_W).
\end{aligned}
$$

2 Quick Review of Continuous Random Variables

References:

- Hoel, Port and Stone, *Introduction to Probability Theory*
- Apostol, Chapter 14

A *(continuous) random variable* X is a variable that can take on any value according to a probability distribution, $\mathbb{P}$:

$$\mathbb{P}(x_l \leq X \leq x_u) = \text{probability that } x \text{ takes on a value in the range } x_l,\, x_u.$$

More generally, we write $\mathbb{P}(A)$ as the probability that an event A will occur (eg, $A = \{x_l \leq X \leq x_u\}$).

Properties:

1. If X is a random variable in the range $[l, u]$ then $\mathbb{P}(l \leq X \leq u) = 1$.
2. If $m \in [l, u]$ then $\mathbb{P}(l \leq X \leq m) = 1 - \mathbb{P}(m \leq X \leq u)$

We characterize a random variable in terms of the *probability density function* (pdf), $p(x)$:

$$\mathbb{P}(x_l \leq X \leq x_u) = \int_{x_l}^{x_u} p(x) dx.$$

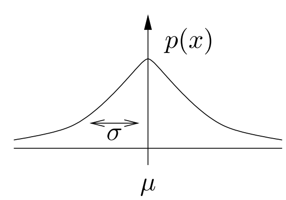
This can be taken as the definition of the pdf. We will sometimes write $p_X(x)$ when we wish to make explicit that the pdf is associated with the random variable X.

Standard pdfs:

Uniform: $p(x) = \frac{1}{U-L}$

Gaussian: (or normal) $p(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}$

μ = mean $\quad$ σ = std dev



There many others as well, but we will generally use Gaussians.

Joint Probability If two random variables are related, we can talk about their joint probability: $\mathbb{P}(A \cap B)$ is the probability that events A and B both occur. For continuous random variables, these can be characterized in terms of a joint probability density function

$$\mathbb{P}(x_l \leq X \leq x_u, y_l \leq Y \leq y_u) = \int_{y_l}^{y_u} \int_{x_l}^{x_u} p(x, y)\, dx\, dy$$

Conditional Probability The conditional probability for an event A given that an event B has occurred, written as $\mathbb{P}(A \mid B)$, is given by

$$\mathbb{P}(A \mid B) = \frac{\mathbb{P}(A \cap B)}{P(B)}$$

where $\mathbb{P}(A \cap B)$ is the probability that both event A and event B occurred. If the events A and B are independent, then $\mathbb{P}(A \mid B) = \mathbb{P}(A)$.

The analog of the probability density function for conditional probability is the conditional probability density function $p(x \mid y)$

$$p(x \mid y) = \begin{cases} \frac{p(x,y)}{p(y)}, & 0 < p(y) < \infty \\ 0 & \text{otherwise.} \end{cases}$$

It follows that

$$p(x, y) = p(x \mid y)\, p(y)$$

and

$$\mathbb{P}(x_l \leq X \leq x_u \mid y) = \int_{x_l}^{x_u} p(x \mid y)\, dx = \frac{\int_{x_l}^{x_u} p(x, y)\, dx}{p(y)}$$

Note that $p(x, y)$ and $p(x \mid y)$ are different density functions.

Remarks:

1. If X and Y are related with conditional probability distribution $p(x \mid y)$ then

$$p(x) = \int_{-\infty}^{\infty} p(x, y)\, dy = \int_{-\infty}^{\infty} p(x \mid y) p(y)\, dy.$$

   This operation is know as the *marginalization* of X.

3 Random Processes

A *random process* is a stochastic system characterized by the *evolution* of a random variable $X(t)$, $t \in [0, T]$. The process is defined in terms of the "correlation" of $X(t_1)$ with $X(t_2)$.

We call $X(t) \in \mathbb{R}^n$ the *state* of the random process and can characterize the state in terms of a time-varying pdf,

$$\mathbb{P}(x_l \leq X(t) \leq x_u) = \int_{x_l}^{x_u} p(x; t)\, dx.$$

Note that the state of a random process is not enough to determine the next state (otherwise it would be a deterministic process).

We can characterize the dynamics of a random process by its statistical characteristics, written in terms of *joint probability* density functions:

$$\mathbb{P}(x_{1l} \leq X(t_1) \leq x_{1u}, x_{2l} \leq X(t_2) \leq x_{2u}) = \int_{x_{2l}}^{x_{2u}} \int_{x_{1l}}^{x_{1u}} p(x_1, x_2; t_1, t_2)\, dx_1\, dx_2.$$

The function $p(x_1, x_2; t_1, t_2)$ is called a *joint probability density function.*

Remarks:

1. In practice, pdf's are not available for most random processes $\Longrightarrow$ mainly useful for analysis.
2. Typically we *assume* a certain pdf (or class of pdfs) as a model, e.g., Gaussian white noise (to be defined later)

Standard definitions

Given a random variable X, we can define various standard measures of the distribution:

$$\begin{aligned}
\mathbb{E}(X(t)) &= \int_{-\infty}^{\infty} x\, p(x,t)\, dx && \text{expectation, } \mu(t) \\
\mathbb{E}(X^2(t)) &= \int_{-\infty}^{\infty} x^2\, p(x,t)\, dx && \text{mean square} \\
\mathbb{E}((X(t)-\mu(t))) &= \int_{-\infty}^{\infty} (x-\mu(t))^2\, p(x,t)\, dx && \text{variance, } \sigma^2(t) \\
\mathbb{E}(X(t_1)X(t_2)) &= \int_{-\infty}^{\infty} x_1 x_2\, p(x_1, x_2; t_1, t_2)\, dx_1 dx_2 && \text{correlation function, } r(t_1, t_2).
\end{aligned}$$

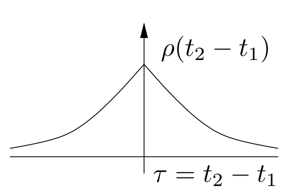
A process is *stationary* if $p(x, t+s) = p(x,t)$ for all s, $p(x_1, x_2; t_1+s, t_2+s) = p(x_1, x_2; t_1, t_2)$, etc. In this case we can write $p(x_1, x_2; \tau)$ for the joint probability distribution, where $\tau = t_2 - t_1$. We will almost always restrict to this case. Also: write $r(t_1, t_2)$ as $r(\tau) = r(t, t+\tau)$.

Example: Consider a first order *Markov process* (also called an Ornstein-Uhlenbeck process) defined by a Gaussian pdf with $\mu = 0$,

$$p(x,t) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2}\frac{x^2}{\sigma^2}},$$

and a correlation function given by

$$r(t_1, t_2) = \frac{Q}{2\omega_0} e^{-\omega_0 |t_2 - t_1|}.$$



This is a stationary process.

Note: we don't usually specify joint probability distributions; we normally use correlation functions.

Property 1 $\mathbb{E}(\alpha X + \beta Y) = \alpha\mathbb{E}(X) + \beta\mathbb{E}(Y)$.

Proof: follows from linearity of integration

Property 2 If X and Y are Gaussian random processes/variables with

$$p(x) = \frac{1}{\sqrt{2\pi\sigma_x^2}} e^{-\frac{1}{2}\left(\frac{x-\mu_x}{\sigma_x}\right)^2} \qquad p(y) = \frac{1}{\sqrt{2\pi\sigma_y^2}} e^{-\frac{1}{2}\left(\frac{y-\mu_y}{\sigma_y}\right)^2}$$

then $X + Y$ is a Gaussian random process/variable with

$$p(x+y) = \frac{1}{\sqrt{2\pi\sigma_z^2}} e^{-\frac{1}{2}\left(\frac{x+y-\mu_z}{\sigma_z}\right)^2},$$

where

$$\mu_z = \mu_x + \mu_y, \qquad \sigma_z^2 = \sigma_x^2 + \sigma_y^2.$$

Proof: Homework. Hint: Use the fact that $p(z \mid y) = p_x(x) = p_x(z - y)$.

Property 3 If X is a Gaussian random process/variable with mean μ and variance σ^2, then αX is Gaussian with mean αX and variance $\alpha^2\sigma^2$.

Proof:

$$\begin{aligned}
\mathbb{P}(x_l \le \alpha X \le x_u) &= \mathbb{P}\left(\frac{x_l}{\alpha} \le X \le \frac{x_u}{\alpha}\right) \\
&= \int_{\frac{x_l}{\alpha}}^{\frac{x_u}{\alpha}} \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} dx \\
&= \int_{x_l}^{x_u} \frac{1}{\alpha\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2}\left(\frac{y/\alpha-\mu}{\sigma}\right)^2} dy \\
&= \int_{x_l}^{x_u} \frac{1}{\sqrt{2\pi\alpha^2\sigma^2}} e^{-\frac{1}{2}\left(\frac{y-\alpha\mu}{\alpha\sigma}\right)^2} dy \quad = \int_{x_l}^{x_u} p(y)\, dy.
\end{aligned}$$

Vector Processes

Suppose

$$X(t) = \begin{bmatrix} X_1(t) \\ \vdots \\ X_n(t) \end{bmatrix}$$

is a *vector* random process. The previous definitions can be extended as

$$\mathbb{E}(X(t)) = \begin{bmatrix} \mathbb{E}(X_1(t)) \\ \vdots \\ \mathbb{E}(X_n(t)) \end{bmatrix},$$

$$\mathbb{E}(X(t_1)X^\mathsf{T}(t_2)) = \begin{bmatrix} \mathbb{E}(X_1(t_1)X_1(t_2)) & \dots & \mathbb{E}(X_1(t_1)X_n(t_2)) \\ & \ddots & \vdots \\ & & \mathbb{E}(X_n(t_1)X_n(t_2)) \end{bmatrix} = R(t_1, t_2).$$

The matrix $R(t_1, t_2) \in \mathbb{R}^{n \times n}$ is called the *correlation matrix* for $X(t) \in \mathbb{R}^n$.

Special case: $R(t, t)$ is called the *covariance matrix*. Defines how elements of x are correlated at time t (with each other). Note that the elements on the diagonal of $R(t, t)$ are the variances of the corresponding variable.

Notation:

- In some communities (eg statistics), the term "cross-covariance" is used to refer to the covariance between two random vectors X and Y, to distinguish this from the covariance of the elements of X with each other. The term "cross-correlation" is sometimes also used.

- NumPy and SciPy have a number of functions to implement covariance and correlation, which mostly match the terminology here:
  - `numpy.cov(Xs)` - returns the sample variance of the vector random variable $X \in \mathbb{R}^n$ where each column of `Xs` represents samples of X.
  - `numpy.cov(Xs, Ys)` - returns the (cross-) covariance of the variables X and Y where `Xs` and `Ys` represent samples of the given random variables.
  - `scipy.correlate(X, Y)` - the "cross-correlation" between two random (1D) sequences. If these sequences came from a random process, this is a single sample approximation of the (discrete-time) correlation function. Use `scipy.correlation_lags` to compute the lag τ.

Example (of how we will use this). Suppose $V(t)$ is a Gaussian random process with $V(t_1)$ and $V(t_2)$ independent for all $t_1 \neq t_2$. Let

$$\dot{X} = AX + FV \qquad X \in \mathbb{R}^n.$$

We can use the above definitions to characterize $X(t) \in \mathbb{R}^n$, now viewed as a random process.

Power Spectral Density

We can also characterize the spectrum of a random process. Let $r(\tau)$ be the correlation function for a random process. We define the *power spectral density function* as

$$S(\omega) = \int_{-\infty}^{\infty} r(\tau)e^{-j\omega\tau}\, d\tau \qquad \text{Fourier transform,}$$
$$r(\tau) = \frac{1}{2\pi}\int_{-\infty}^{\infty} S(\omega)e^{j\omega\tau}\, d\tau \quad \text{Inverse Fourier transform.}$$

Definition 1. A process is *white noise* if $\mathbb{E}(X(t)) = 0$ and $S(\omega) = W = \text{constant}$ for all ω. If $X(t) \in \mathbb{R}^n$ (a random vector), then $W \in \mathbb{R}^{n\times n}$.

Properties:

1. $r(\tau) = \frac{1}{2\pi}\int_{-\infty}^{\infty} S(\omega)e^{j\omega\tau}\, d\tau = W\delta(\tau)$, where $\delta(\tau)$ is the unit impulse.

   Proof: If $\tau \neq 0$ then

   $$r(\tau) = \frac{1}{2\pi}\int_{-\infty}^{\infty} W(\cos(\omega\tau) + j\sin(\omega\tau)\, d\tau = 0.$$

   If $\tau = 0$ then $r(\tau) = \infty$. Can show that

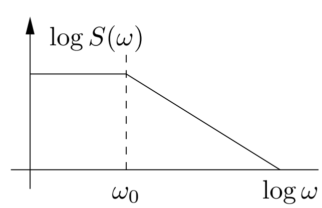
   $$\int_{-\epsilon}^{\epsilon}\int_{-\infty}^{\infty}(\cdots)\, d\omega d\tau = W.$$

2. $r(0) = \mathbb{E}(x^2(t)) = \infty \implies$ idealization; never see this in practice.
3. Typically we use white noise as an idealized input to a system and characterize the output of the system (similar to the impulse response in deterministic systems).

Example First order Markov process

$$r(\tau) = \frac{Q}{2\omega_0}e^{-\omega_0(\tau)} \qquad p(x) = \frac{1}{\sqrt{2\pi\sigma^2}}e^{-\frac{x^2}{2\sigma^2}}$$

$$\begin{aligned} S(\omega) &= \int_{-\infty}^{\infty}\frac{Q}{2\omega_0}e^{-\omega|\tau|}e^{-j\omega\tau}\, d\tau \\ &= \int_{-\infty}^{0}\frac{Q}{2\omega_0}e^{(\omega-j\omega)\tau}\, d\tau + \int_{0}^{\infty}\frac{Q}{2\omega_0}e^{(-\omega-j\omega)\tau}\, d\tau \\ &= \frac{Q}{\omega^2+\omega_0^2} \end{aligned}$$



Be careful: $S(\omega)$ is *not* a transfer function. $S(\omega)$ is always real.

4 Linear Stochastic Systems

We now consider the problem of how to compute the response of a linear system to a random variable. We assume we have a linear system described either in state space or as a transfer function:

$$\begin{aligned} \dot{X} &= AX + FV, \\ Y &= CX, \end{aligned} \qquad H(s) = C(sI - A)^{-1}F.$$

Given an input V which is itself a random process with mean $\mu(t)$, variance $\sigma^2(t)$ and correlation $r(t)$, what is the description of the random process Y?

Spectral Response

Let V be a (scalar) Gaussian white noise process, with zero mean and covariance Q:

$$r(\tau) = Q\delta(\tau).$$

We can write the output of the system in terms of the convolution integral

$$Y(t) = \int_0^t h(t-\tau)V(\tau)\,d\tau \qquad h(t-\tau) = Ce^{A(t-\tau)}F \quad \text{(impulse response)}.$$

We now compute the statistics of the output, starting with the mean

$$\begin{aligned} \mathbb{E}(Y) &= \mathbb{E}(\int_0^t h(t-\eta)V(\eta)\,d\eta) \\ &= \int_0^t h(t-\eta)\mathbb{E}(V(\eta))\,d\eta = 0 \quad \text{(zero mean)}. \end{aligned}$$

Note here that we have relied on the linearity of the convolution integral to pull the expectation inside the integral.

We can compute the covariance of the output by computing the correlation $r(\tau)$ and setting $\sigma^2 = r(0)$. The correlation function for y is

$$\begin{aligned} r_Y(t_1, t_2) = \mathbb{E}(Y(t_1)Y(t_2)) &= \mathbb{E}\left(\int_0^{t_1} h(t_1-\eta)V(\eta)\,d\eta \cdot \int_0^{t_2} h(t_2-\xi)V(\xi)\,d\xi\right) \\ &= \mathbb{E}\left(\int_0^{t_1}\int_0^{t_2} h(t_1-\eta)V(\eta)V(\xi)h(t_2-\xi)\,d\eta d\xi\right). \end{aligned}$$

Once again linearity allows us to exchange expecation and integration

$$\begin{aligned} r_y(t_1,t_2) &= \int_0^{t_1}\int_0^{t_2} h(t_1-\eta)\mathbb{E}(V(\eta)V(\xi))h(t_2-\xi)\,d\eta d\xi \\ &= \int_0^{t_1}\int_0^{t_2} h(t_2-\eta)Q\delta(\eta-\xi)h(t_2-\xi)\,d\eta d\xi \\ &= \int_0^{t_1} h(t_1-\eta)Qh(t_2-\eta)\,d\eta. \end{aligned}$$

Now let $\tau = t_2 - t_1$ and write

$$\begin{aligned} r_y(\tau) = r_y(t,t+\tau) &= \int_0^{t} h(t-\eta)Qh(t+\tau-\eta)\,d\eta \\ &= \int_0^{t} h(\xi)Qh(\xi+\tau)\,d\eta \qquad \text{(setting } \xi = t-\eta). \end{aligned}$$

Finally, we let $t \to \infty$ (steady state):

$$\lim_{t\to\infty} r_Y(t,t+\tau) = r_Y(\tau) = \int_0^{\infty} h(\xi)Qh(\xi+\tau)d\xi.$$

If this integral exists, then we can compute the second order statistics for the output Y.

We now compute the spectral density function corresponding to Y:

$$\begin{aligned} S_Y(\omega) &= \int_{-\infty}^{\infty}\left[\int_0^{\infty} h(\xi)Qh(\xi+\tau)d\xi\right] e^{-j\omega\tau}\,d\tau \\ &= \int_0^{\infty} h(\xi)Q\left[\int_{-\infty}^{\infty} h(\xi+\tau)e^{-j\omega\tau}\,d\tau\right]d\xi \\ &= \int_0^{\infty} h(\xi)Q\left[\int_0^{\infty} h(\lambda)e^{-j\omega(\lambda-\xi)}\,d\lambda\right]d\xi \\ &= \int_0^{\infty} h(\xi)e^{j\omega\xi}\,d\xi\cdot QH(j\omega) = H(-j\omega)Q_V H(j\omega). \end{aligned}$$

This is then the response of a linear system to white noise.

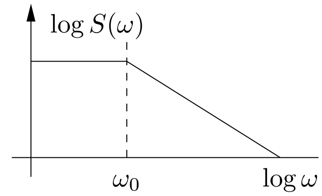
Composition:

$$S_Y(\omega) = H_2(-j\omega)H_1(-j\omega)Q_V H_1(j\omega)H_2(j\omega).$$

Note that in the frequency domain, we get *multiplication* of frequency responses (same as composition transfer functions)

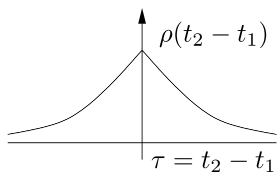
Example: first order response

$$\begin{aligned} H(j\omega) &= \frac{1}{s+\omega_0} \qquad \Longrightarrow \\ S_Y(\omega) &= \frac{1}{-j\omega+\omega_0}\cdot Q\cdot\frac{1}{j\omega+\omega_0} = \frac{Q}{\omega^2+\omega_0^2}. \end{aligned}$$



We can also compute the correlation function for a random process Y:

$$r_Y(\tau) = \int_0^\infty h(\xi) Q h(\xi+\tau) d\xi = \frac{Q}{2\omega_0} e^{-\omega|\tau|}$$



- Shows that $Y(t)$ is correlated to $Y(t+\tau)$ even though input is white (uncorrelated)
- Correlation drops off as τ increases

Spectral factorization

We often want to find a Q and H such that we match the statistics of measured noise. Eg, given $S(\omega)$, find $Q > 0$ and $H(s)$ such that $S(\omega) = H(-j\omega) Q H(j\omega)$.

Exercise Find a constant matrix A and vectors F and C such that for

$$\dot{X} = AX + FW, \; Y = CX$$

the power spectrum of Y is given by

$$S(\omega) = \frac{1+\omega^2}{(1-7\omega^2)^2 + 1}$$

Describe the sense in which your answer is unique.

State Space Computations

We now consider the computation of random processes associated with state space representations of a system

$$\begin{aligned} \dot{X} &= AX + FV, \\ Y &= CX. \end{aligned} \qquad (1)$$

Q: What is $r_Y(\tau)$ in terms of A, F, C and $r_v(\tau)$?

We will consider the general case where $V \in \mathbb{R}^m$, $Y \in \mathbb{R}^p$ and $r \to R$.

Define the *state transition matrix* $\Phi(t, t_0) = e^{A(t-t_0)}$ so that the solution of system (1) is given by

$$x(t) = \Phi(t, t_0) x(t_0) + \int_{t_0}^t \Phi(t, \lambda) F v(\lambda) d\lambda.$$

Claim Let $\mathbb{E}(X(t_0) X^\mathsf{T}(t_0)) = P(t_0)$ and V be white noise with $\mathbb{E}(V(\lambda) V^\mathsf{T}(\xi)) = Q_v \delta(\lambda - \xi)$. Then the correlation matrix for X is given by

$$R_X(t_1, t_2) = P(t_1) \Phi^\mathsf{T}(t_2, t_1),$$

where

$$P(t) = \Phi(t,t_0)P(t_0)\Phi^\mathsf{T}(t,t_0) + \int_{t_0}^{t} \Phi(t,\lambda)FQ_VF^\mathsf{T}\Phi^\mathsf{T}(t,\lambda)\,d\lambda.$$

Proof.

$$\begin{aligned}
\mathbb{E}(X(t_1)X^\mathsf{T}(t_2)) &= \mathbb{E}\Big(\Phi(t_1,0)X(0)X^\mathsf{T}(0)\Phi^\mathsf{T}(t_2,0) + \text{cross terms} \\
&\qquad + \int_0^{t_1} \Phi(t_1,\xi)FV(\xi)d\xi \int_0^{t_2} V^\mathsf{T}(\lambda)F^\mathsf{T}\Phi(t_2,\lambda)d\lambda\Big) \\
&= \Phi(t_1,0)\mathbb{E}(X(0)X^\mathsf{T}(0))\Phi(t_1,0) \\
&\qquad + \int_0^{t_1}\int_0^{t_2} \Phi(t_1,\xi)F\mathbb{E}(V(\xi)V^\mathsf{T}(\lambda))F^\mathsf{T}\Phi(t_2,\lambda)d\xi d\lambda \\
&= \Phi(t_1,0)P(0)\Phi^\mathsf{T}(t_2,0) + \int_0^{t_1} \Phi(t_2,\lambda)FQ_V(\lambda)F^\mathsf{T}\Phi(t_2,\lambda)d\lambda.
\end{aligned}$$

Now use the fact that $\Phi(s,0) = \Phi(s,t)\Phi(t,0)$ (and similar relations) to obtain

$$R_X(t_1,t_2) = P(t_1)\Phi^\mathsf{T}(t_2,t_1),$$

where

$$P(t) = \Phi(t,0)P(0)\Phi^\mathsf{T}(t,0) + \int_0^\mathsf{T} \Phi(t,\lambda)FQ_VF^\mathsf{T}(\lambda)\Phi^\mathsf{T}(t,\lambda)d\lambda.$$

Finally, differentiate to obtain

$$\dot{P}(t) = AP + PA^\mathsf{T} + FQ_VF, \qquad P(0) \;= P_0$$

(see Friedland for details). □

For time-invariant systems, can show that $P(t)$ approaches a *constant*, so that

$$R_X(\tau) = R_X(t,t+\tau) = Pe^{A\tau}$$

and P satisfies the algebraic equation

$$AP + PA^\mathsf{T} + FQ_VF^\mathsf{T} = 0 \qquad P > 0.$$

This is called the *Lyapunov equation* and can be solved in Python using the function `control.lyap(A, Q)`, where $Q = FR_VF^\mathsf{T}$.

Remarks:

1. For the single state case, can use this to get the same answer as before: $R_X(\tau)$ dies off exponentially in $\tau \implies$ decreasing correlation between $X(t)$ and $X(t+\tau)$ as τ increases.

2. Can also compute the correlation matrix for the output:

$$
\begin{aligned}
&Y = CX \qquad \Longrightarrow \qquad Y(t)Y^{\mathsf{T}}(t) = CX(t)X^{\mathsf{T}}(t)C^{\mathsf{T}} \\
&R_Y(\tau) = \mathbb{E}(Y(t)Y^{\mathsf{T}}(t+\tau)) = CR_X(\tau)C^{\mathsf{T}} \\
&R_Y(0) = CPC^{\mathsf{T}}.
\end{aligned}
$$

Summary of random processes:

RMM: Check

$$
\begin{array}{ccc}
p(v) = \dfrac{1}{\sqrt{2\pi Q_V}} e^{-\frac{x^2}{2Q_V}} & & p(y) = \dfrac{1}{\sqrt{2\pi R_Y}} e^{-\frac{x^2}{2R_Y}} \\
 & V \longrightarrow \boxed{\;H\;} \longrightarrow Y & \\
S_V(\omega) = Q_V & & S_Y(\omega) = H(-j\omega)Q_V H(j\omega) \\
 & & \\
r_V(\tau) = Q_V\delta(\tau) & \begin{aligned} \dot{X} &= AX + FV \\ Y &= CX \end{aligned} & \begin{aligned} r_Y(\tau) = R_Y(\tau) = CPe^{-A\tau}C^{\mathsf{T}} \\ AP + PA^{\mathsf{T}} + FQ_VF^{\mathsf{T}} = 0 \end{aligned}
\end{array}
$$