System Identification

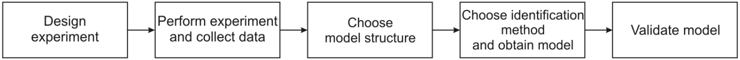
Control Engineering EN, 3rd year B.Sc.
Technical University of Cluj-Napoca
Romania

Lecturer: Lucian Buşoniu

Part VI

Input signals

Motivation



Choosing inputs is the core of experiment design

All identification methods require inputs to satisfy certain conditions, for example:

- Transient analysis requires step or impulse inputs
- Correlation analysis preferably works with white-noise input
- ARX requires “sufficiently informative” inputs

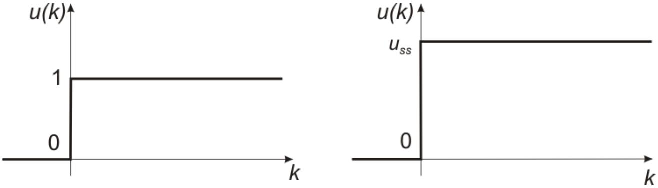
Plan

In this part we:

- Revisit some types of input signals that were already used
- Describe a few new types of input signals
- Discuss choices and properties of input signals important for system identification
- Characterize the signals discussed using the properties introduced

Table of contents

Step input



Left: Unit step:

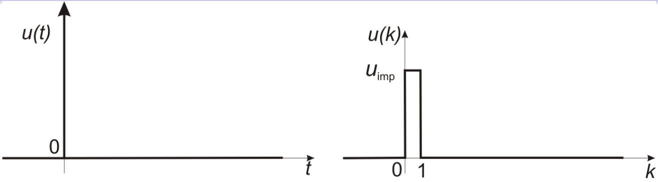
$$u(k) = \begin{cases} 0 & k < 0 \\ 1 & k \geq 0 \end{cases}$$

Right: Step of arbitrary magnitude:

$$u(k) = \begin{cases} 0 & k < 0 \\ u_{ss} & k \geq 0 \end{cases}$$

Remark: These are discrete-time reformulations of the continuous-time variants.

Impulse input



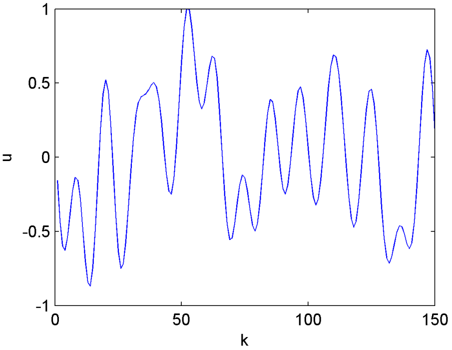
In discrete-time, we cannot freely approximate the ideal impulse (left), since the signal can only change values at the sampling instants.

Right: Discrete-time impulse realization:

$$u(k) = \begin{cases} u_{\text{imp}} & k = 0 \\ 0 & \text{otherwise} \end{cases}$$

- When $u_{\text{imp}} = \frac{1}{T_s}$, the integral of the signal is 1 and we get an approximation of the continuous-time impulse.
- When $u_{\text{imp}} = 1$ (e.g. in correlation analysis), we get a “unit” discrete-time impulse.

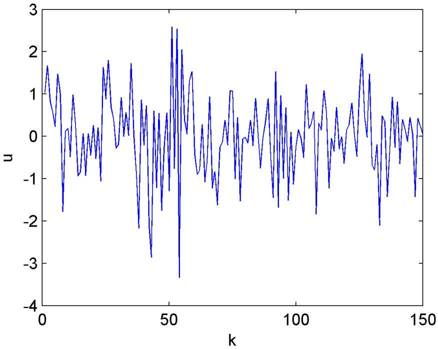
Sum of sines



$$u(k) = \sum_{j=1}^{m} a_j sin(\omega_j k + \varphi_j)$$

- a_j: amplitudes of the m component sines
- ω_j: frequencies, $0 \leq \omega_1 < \omega_2 < \ldots < \omega_m \leq \pi$
- φ_j: phases

White noise

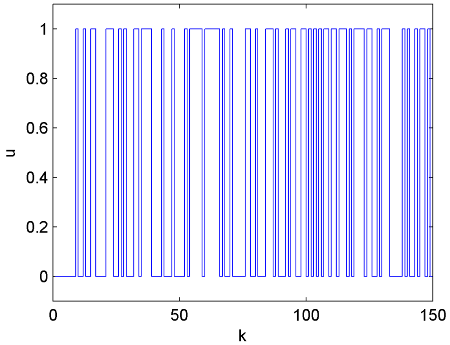


Recall zero-mean white noise: mean 0, different steps uncorrelated.

In the figure, values were independently drawn from a zero-mean Gaussian distribution.

Table of contents

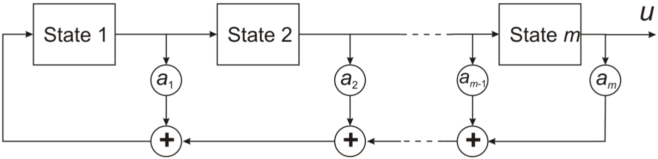
Pseudo-random binary sequence (PRBS)



A signal that switches between two discrete values, generated with a specific algorithm.

Interesting because it approximates white noise, and so it inherits some of the useful properties of white noise (formalized later).

PRBS generator



PRBS can be generated with a linear shift feedback register as in the figure. All signals and coefficients are binary (the states are bits).

At each discrete step k:

- State x_i transfers to state x_{i+1}.
- State x_1 is set to the modulo-two addition of states on the feedback path (if $a_i = 1$ then x_i is added, if $a_i = 0$ then it is not).
- Output $u(k)$ is collected at state x_m.

Remark: such a feedback register is easily implemented in hardware.

Modulo-two addition

Formula/truth table of modulo-two addition:

$$p \oplus q = \begin{cases} 0 & \text{if } p = 0, q = 0 \\ 1 & \text{if } p = 0, q = 1 \\ 1 & \text{if } p = 1, q = 0 \\ 0 & \text{if } p = 1, q = 1 \end{cases}$$

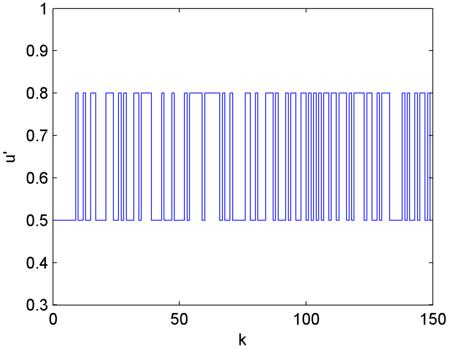
...also known as XOR (eXclusive OR)

Arbitrary-valued PRBS

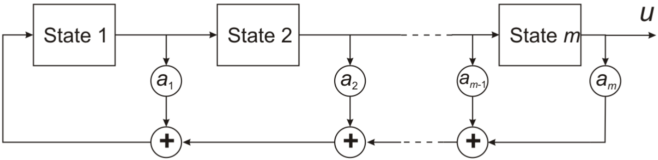
To obtain a signal $u'(k)$ taking values b, c instead of $0, 1$, shift & scale the original signal $u(k)$:

$$u'(k) = b + (c - b)u(k)$$

Example for $b = 0.5$, $c = 0.8$:



State space representation



$$
\begin{aligned}
x_1(k+1) &= a_1 x_1(k) \oplus a_2 x_2(k) \oplus \cdots \oplus a_m x_m(k) \\
x_2(k+1) &= x_1(k) \\
&\vdots \\
x_m(k+1) &= x_{m-1}(k) \\
u(k) &= x_m(k)
\end{aligned}
$$

$x(k) = [x_1(k), \ldots, x_m(k)]^\top$ compactly denotes the state vector of m variables (bits)

State space representation: matrix form

$$x(k+1) = \begin{bmatrix} a_1 & a_2 & \dots & a_{m-1} & a_m \\ 1 & 0 & \dots & 0 & 0 \\ 0 & 1 & \dots & 0 & 0 \\ \vdots & & & & \\ 0 & 0 & \dots & 1 & 0 \end{bmatrix} \otimes x(k) =: A \otimes x(k)$$

$$u(k) = [0\ 0\ \dots\ 0\ 1]x(k) =: Cx(k)$$

where $\otimes$ symbolically indicates that the additions in the matrix product are performed modulo 2.

Period of PRBS

- The PRBS algorithm is deterministic, so the current state $x(k)$ fully determines the future states and outputs

$\Rightarrow$ Period (number of steps until sequence repeats) at most 2^m

- The identically zero state is undesirable, as the future sequence would always remain 0

$\Rightarrow$ Maximum practical period is $P = 2^m - 1$

A PRBS with period $P = 2^m - 1$ is called maximum-length PRBS.

Such PRBS have interesting characteristics, so they are preferred in practice.

Maximum-length PRBS

The period is determined by the feedback coefficients a_i.

The following coefficients must be 1 to achieve maximum length (all others 0):

m	Max period $2^m - 1$	Coefficients equal to 1
3	7	a_1, a_3
4	15	a_1, a_4
5	31	a_2, a_5
6	63	a_1, a_6
7	127	a_1, a_7
8	255	a_1, a_2, a_7, a_8
9	511	a_4, a_9
10	1023	a_3, a_{10}

Other working combinations of coefficients exist, and coefficients for larger m can be found in the literature.

Matlab function

```
u = idinput(N, type, [], [b, c]);
```

Arguments:

1. `N`: signal length (number of discrete steps).
2. `type`: signal type, a string. Relevant for us: `'prbs'` for PRBS, `'rgs'` for white Gaussian noise, `'sin'` for multisine.
3. Third argument: the frequency band of the inputs (can be left at its default, empty matrix).
4. `[b, c]`: the range (lower and upper limits) of the signal.

Remark: `N` can be configured to generate multiple-input signals (see the Matlab documentation for details).

Table of contents

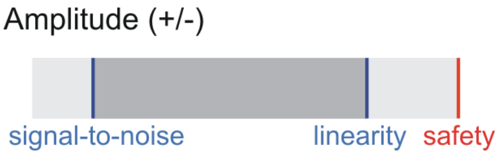
Choice of input shape

Some identification methods require specific types of inputs:

- Transient analysis requires step or impulse inputs.
- Correlation analysis preferably works with white-noise input.

Rule of thumb: input shapes, including characteristics like amplitude, should be chosen to be representative for the typical operation of the system

Choice of input amplitude



- Range of allowed inputs typically constrained by system operator, due to safety or cost concerns
- Even if allowed, overly large inputs may take the system out of its zone of linearity and lead to poor performance of linear identification
- But too small inputs will lead to signals dominated by noise and disturbance

Choice of sampling interval

For nearly all methods, we work in discrete time so we must choose a sampling interval T_s

- Too large intervals will not model the relevant dynamics of the system. Initial idea: 10% of the smallest time constant
- Too small intervals will lead to overly large effects of noise and disturbance
- When in doubt, take T_s smaller

Due to Nyquist-Shannon, we know that signals cannot be recovered above frequency $1/(2T_s)$, so to mitigate noise and other effects it is useful to pass the outputs (and inputs, if measured) through a low-pass filter that eliminates higher frequencies

Mean and covariance

Given a random signal $u(k)$, its mean and covariance are defined:

$$\mu = \mathrm{E}\left\{u(k)\right\}$$
$$r_u(\tau) = \mathrm{E}\left\{[u(k+\tau) - \mu][u(k) - \mu]\right\}$$

Recall:

- Mean and variance of random variables
- Related covariance function $r_u(\tau)$ in correlation analysis (there we assumed μ was already 0)

Mean and covariance: deterministic signal

When the signal is deterministic (e.g. PRBS), the mean and covariance are redefined as:

$$\mu = \lim_{N \to \infty} \frac{1}{N} \sum_{k=1}^{N} u(k)$$

$$r_u(\tau) = \lim_{N \to \infty} \frac{1}{N} \sum_{k=1}^{N} [u(k+\tau) - \mu][u(k) - \mu]$$

Note: $\lim_{N \to \infty} \frac{1}{N} \sum_{k=1}^{N} \cdot$ is the same as $\mathrm{E}\left\{\cdot\right\}$ for a (well-behaved) random signal.

Generalization to vector signals $u(k) \in \mathbb{R}^{nu}$: interpret the sums elementwise, replace $[u(k+\tau) - \mu][u(k) - \mu]$ by $[u(k+\tau) - \mu][u(k) - \mu]^\top$, an $nu \times nu$ covariance matrix.

Handling nonzero means

- Correlation analysis requires zero-mean signals
- But even other methods (like ARX or more general prediction error methods) may work better when the means are removed

The means can be removed from the signal, and then possibly modeled separately

(see Söderström and Stoica, Chapter 12 for more details)

Persistent excitation

Even methods that do not fix the input shape make requirements on the inputs: e.g. for ARX we required that $u(k)$ is “sufficiently informative”, without making that property formal

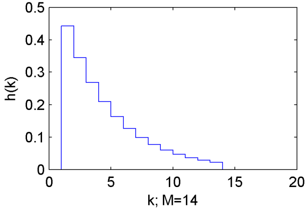
This condition can be precisely stated in terms of a property called persistence of excitation

Persistent excitation: Motivating example

We develop an *idealized* version of correlation analysis. This is only an intermediate motivating step, and the property is useful in many identification algorithms.

Finite impulse response (FIR) model:

$$y(k) = \sum_{j=0}^{M-1} h(j)u(k-j) + v(k)$$



Correlation analysis: Covariances

Assuming $u(k)$, $y(k)$ are zero-mean, so the means do not need to be subtracted:

$$r_u(\tau) = \lim_{N\to\infty} \frac{1}{N} \sum_{k=1}^{N} u(k+\tau)u(k)$$

$$r_{yu}(\tau) = \lim_{N\to\infty} \frac{1}{N} \sum_{k=1}^{N} y(k+\tau)u(k)$$

In practice covariances must be estimated from finite datasets, but here we work with their ideal values (since this is only a motivating example, which we do not actually implement).

Correlation analysis: Identifying the FIR

Taking M equations to find the FIR parameters, we have:

$$\begin{bmatrix} r_{yu}(0) \\ r_{yu}(1) \\ \vdots \\ r_{yu}(M-1) \end{bmatrix} = \boxed{\begin{bmatrix} r_u(0) & r_u(1) & \dots & r_u(M-1) \\ r_u(1) & r_u(0) & \dots & r_u(M-2) \\ \vdots & & & \\ r_u(M-1) & r_u(M-2) & \dots & r_u(0) \end{bmatrix}} \cdot \begin{bmatrix} h(0) \\ h(1) \\ \vdots \\ h(M-1) \end{bmatrix}$$

We are allowed to take a square system (number of equations equal to number of parameters) because we are in the idealized, noise-free case, so overfitting is not a concern.

Denote the matrix in the equation by $R_u(M)$, the covariance matrix of the input.

Persistent excitation: formal definition

Definition

A signal $u(k)$ is persistently exciting (PE) of order n if $R_u(n)$ is positive definite.

A matrix $A \in \mathbb{R}^{n\times n}$ is positive definite if $h^\top Ah > 0$ for any nonzero vector $h \in \mathbb{R}^n$. Note that A must be nonsingular.

Examples:

- $\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ is positive definite. Denote $h = \begin{bmatrix} a \\ b \end{bmatrix}$, then $h^\top Ah = a^2 + b^2$.
- $\begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix}$ is not positive definite. Counterexample: $h = \begin{bmatrix} a \\ -a \end{bmatrix}$, $h^\top Ah = -2a^2$.

PE in correlation analysis

If the order of PE is M, then $R_u(M)$ is positive definite, hence invertible and the linear system from correlation analysis can be solved to find an FIR of length M.

So an order M of PE means that an FIR model of length M is identifiable (M parameters can be found).

General role of PE

Beyond FIR, PE plays a role in *all* parametric system identification methods, including ARX and methods still to be discussed, like prediction error methods and instrumental variable techniques.

A large enough order of PE is required to properly identify the parameters.

Typically, the required order is a multiple of (usually twice) the number of parameters n that must be estimated.

Covariance alternatives

In the sequel we will always use the following covariance definition:

$$\boxed{r_u(\tau) = \lim_{N\to\infty} \frac{1}{N} \sum_{k=1}^{N} u(k+\tau)u(k)}$$

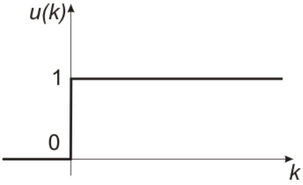
disregarding that this is the true covariance only when $u(k)$ is zero-mean. The resulting function is still useful, and more convenient to compute.

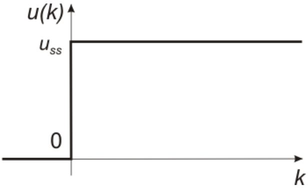
When applying the PE condition for nonzero-mean signals, the simplified definition above will lead to an order of PE larger by 1 than the order of PE obtained with the true covariance function:

$$r_u(\tau) = \lim_{N\to\infty} \frac{1}{N} \sum_{k=1}^{N} [u(k+\tau) - \mu][u(k) - \mu]$$

Table of contents

Step input





Take the more general, non-unit step:

$$u(k) = \begin{cases} 0 & k < 0 \\ u_{ss} & k \geq 0 \end{cases}$$

Step input: Mean and covariance

Mean and covariance:

$$\mu = \lim_{N\to\infty} \frac{1}{N} \sum_{k=0}^{N-1} u(k) = u_{ss}$$

$$r_u(\tau) = \lim_{N\to\infty} \frac{1}{N} \sum_{k=0}^{N-1} u(k+\tau)u(k) = u_{ss}^2$$

Note the signal starts from $k = 0$, so the summation is modified (unimportant to the final result).

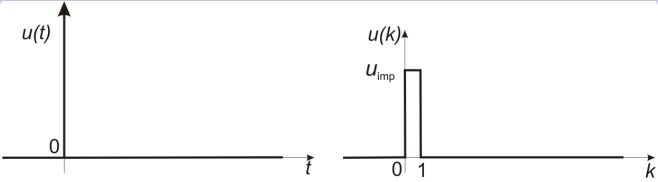
Step input: Order of PE

Covariance matrix:

$$R_u(n) = \begin{bmatrix} r_u(0) & r_u(1) & \dots & r_u(n-1) \\ r_u(1) & r_u(0) & \dots & r_u(n-2) \\ \vdots & & & \\ r_u(n-1) & r_u(n-2) & \dots & r_u(0) \end{bmatrix} = \begin{bmatrix} u_{ss}^2 & u_{ss}^2 & \dots & u_{ss}^2 \\ u_{ss}^2 & u_{ss}^2 & \dots & u_{ss}^2 \\ \vdots & & & \\ u_{ss}^2 & u_{ss}^2 & \dots & u_{ss}^2 \end{bmatrix}$$

This matrix has rank 1, so a step input is PE of order 1.

Impulse input



Recall discrete-time realization:

$$u(k) = \begin{cases} \frac{1}{T_s} & k = 0 \\ 0 & \text{otherwise} \end{cases}$$

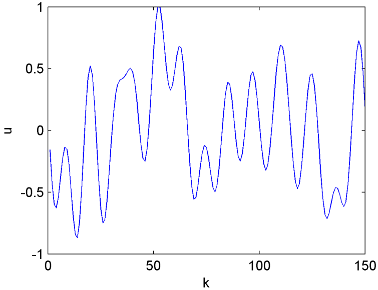
Mean and covariance:

$$\mu = \lim_{N\to\infty} \frac{1}{N} \sum_{k=0}^{N-1} u(k) = 0$$

$$r_u(\tau) = \lim_{N\to\infty} \frac{1}{N} \sum_{k=0}^{N-1} u(k+\tau)u(k) = 0$$

$\Rightarrow$ $R_u(n)$ matrix of zeros, the impulse is not PE of any order.

Sum of sines



$$u(k) = \sum_{j=1}^{m} a_j sin(\omega_j k + \varphi_j), \quad 0 \leq \omega_1 < \omega_2 < \ldots < \omega_n \leq \pi$$

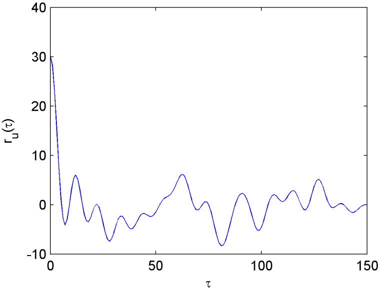
Mean and covariance:

$$\mu = \begin{cases} a_1 sin(\varphi_1) & \text{if } \omega_1 = 0 \\ 0 & \text{otherwise} \end{cases}$$

$$r_u(\tau) = \sum_{j=1}^{m-1} \frac{a_j^2}{2} \cos(\omega_j \tau) + \begin{cases} a_m^2 \sin^2 \varphi_m & \text{if } \omega_m = \pi \\ \frac{a_m^2}{2} \cos(\omega_m \tau) & \text{otherwise} \end{cases}$$

Sum of sines (continued)

For the multisine exemplified before, the covariance function is:



A multisine having m components is PE of order n with:

$$n = \begin{cases} 2m & \text{if } \omega_1 \neq 0, \omega_m \neq \pi \\ 2m-1 & \text{if } \omega_1 = 0 \text{ or } \omega_m = \pi \\ 2m-2 & \text{if } \omega_1 = 0 \text{ and } \omega_m = \pi \end{cases}$$

White noise: Mean and covariance

Take a zero-mean white noise signal of variance σ^2, e.g. drawn from a Gaussian:

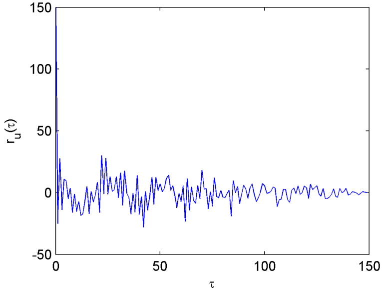
$$u(k) \sim \mathcal{N}(0, \sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{x^2}{2\sigma^2}\right)$$

Then, by definition:

$$\mu = 0$$
$$r_u(\tau) = \begin{cases} \sigma^2 & \text{if } \tau = 0 \\ 0 & \text{otherwise} \end{cases}$$

White noise: Covariance example

Covariance function of white noise signal exemplified before:



White noise: Order of PE

Covariance matrix:

$$R_u(n) = \begin{bmatrix} r_u(0) & r_u(1) & \dots & r_u(n-1) \\ r_u(1) & r_u(0) & \dots & r_u(n-2) \\ \vdots & & & \\ r_u(n-1) & r_u(n-2) & \dots & r_u(0) \end{bmatrix}$$

$$= \begin{bmatrix} \sigma^2 & 0 & \dots & 0 \\ 0 & \sigma^2 & \dots & 0 \\ \vdots & & & \\ 0 & 0 & \dots & \sigma^2 \end{bmatrix} = \sigma^2 I_n$$

where I_n = the identity matrix, positive definite.

$\Rightarrow$ for any n, $R_u(n)$ positive definite — white noise is PE of any order.

Question

Given the information above, why does correlation analysis prefer white noise to other input signals, in order to identify the FIR?

PRBS: Mean

Consider a $0, 1$-valued, maximum-length PRBS with m bits: $P = 2^m - 1$, a large number.

Then its state $x(k)$ will contain all possible binary values with m digits except 0.

Signal $u(k)$ is the last position of $x(k)$, which takes value 1 a number of 2^{m-1} times, and value 0 a number of $2^{m-1} - 1$ times.

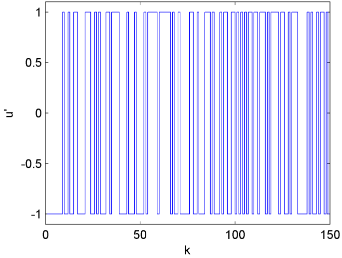
$\Rightarrow$ Mean value:

$$\mu = \frac{1}{P} \sum_{k=1}^{P} u(k) = \frac{1}{P} 2^{m-1} = \frac{(P+1)/2}{P} = \frac{1}{2} + \frac{1}{2P} \approx \frac{1}{2}$$

where the approximation holds for large P.

PRBS: Covariance

Consider a zero-mean PRBS, scaled between $-b$ and b:



$$u'(k) = -b + 2bu(k)$$

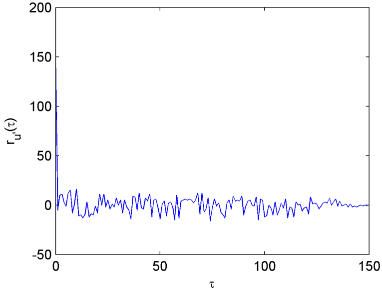
Then:

$$\mu = -b + 2b(\frac{1}{2} + \frac{1}{2P}) = \frac{b}{P} \approx 0$$

$$r_u(\tau) = \begin{cases} 1 - \frac{1}{P^2} \approx 1 & \text{if } \tau = 0 \\ -\frac{1}{P} - \frac{1}{P^2} \approx -\frac{1}{P} \approx 0 & \text{otherwise} \end{cases}$$

PRBS: Covariance example

Covariance function of the zero-mean PRBS above:



So, PRBS behaves similarly to white noise (similar covariance function). Combined with the ease of generating it, this property makes PRBS very useful in system identification.

PRBS: Order of PE

A maximum-length PRBS is PE of exactly order P, the period (and not larger).

Exercise

Using the approximate formula that ignores the terms $\frac{1}{P^2}$:

$$r_u(\tau) \approx \begin{cases} 1 & \text{if } \tau = 0 \\ -\frac{1}{P} & \text{otherwise} \end{cases}$$

take a small value of $P \geq 2$ and show that the PRBS is exactly of PE order P.

Hint: construct $R_u(n)$ for $n = P$ and show that it is rank P, then for $n > P$ and show it is *still* only of rank P. This can be done by showing that columns $P + 1, P + 2, \ldots$ are linear combinations of the first P columns.