# SSY230 Learning dynamical systems using system identification

Hand In 1

Alfred Aronsson

April 26, 2024

# 1 Introduction

First hand in task in SSY230 focused on asymptotic Gaussian distribution of the estimated parameters of the least square solution.

# Problem definition

Assume that the noise sequence $\{e(t)\}_{t=1}^{N}$ consists of independent normally distributed components $e(t) \sim \mathcal{N}(0, \sigma^2)$. Show that the least-squares estimate $\hat{\theta}$ of the parameters of the model $y(t) = \phi^\top(t)\theta + e(t)$ is asymptotically normally distributed. Make necessary assumptions on $\phi(t)$.

*Note:* assume that mean values converge to expectation, as described in S&S Lemma B.1 and Lemma B.3.

# Lemma B.1

Assume that $x(t)$ is a discrete time stationary process with finite variance. If the covariance function $r_x(\tau) \to 0$ as $|\tau| \to \infty$ then

$$\frac{1}{N}\sum_{t=1}^{N} x(t) \to E\{x(t)\} \quad (N \to \infty)$$

with probability one and in mean square.

*Proof.* See, for example, Gnedenko (1963).

# Lemma B.3

Consider

$$X_N = \frac{1}{\sqrt{N}}\sum_{t=1}^{N} z(t)$$

where $z(t)$ is a (vector-valued) zero mean stationary process given by

$$z(t) = \phi(t)v(t)$$

In (B.8), $\phi(t)$ is a matrix and $v(t)$ a vector. The entries of $\phi(t)$ and $v(t)$ are stationary, possibly correlated, ARMA processes with zero means and underlying white noise sequences with finite fourth-order moments. The elements of $\phi(t)$ may also contain a bounded deterministic term.

Assume that the limit

$$P = \lim_{N\to\infty} E X_N X_N^\top$$

exists and is nonsingular. Then $X_N$ is asymptotically Gaussian distributed,

$$X_N \xrightarrow{\text{dist}} \mathcal{N}(0, P)$$

*Proof.* See Ljung (1977c).

# Slutsky's Theorem

Let $\{X_n\}$ be a sequence of random variables that converges in distribution to a random variable $X$ (denoted as $X_n \xrightarrow{d} X$), and let $\{Y_n\}$ be a sequence of random variables that converges in probability to a constant $c$ (denoted as $Y_n \xrightarrow{p} c$). Then the following holds:

1. $X_n + Y_n \xrightarrow{d} X + c$ *as* $n \to \infty$.
2. $X_n Y_n \xrightarrow{d} cX$ *as* $n \to \infty$.

where $\xrightarrow{d}$ denotes convergence in distribution, and $\xrightarrow{p}$ denotes convergence in probability.

# Assumptions

Following from Lemma B.1 we assume that the covariance function $r_{\phi\phi^\top}(\tau)$ goes to zero as $\tau$ goes to infinity. We also assume that $\phi\phi^\top$ has finite variance and is stationary.

Following from Lemma B.3 we assume that $\phi\phi^\top$ is full rank and that $\phi e$ is a stationary ARMA process with zero means with underlying white noise sequences with finite fourth order moments.

## Solution

We can start at the solution of the least square problem,

$$\hat{\theta} = (\phi^T \phi)^{-1} \phi^\top Y \tag{1}$$

where we know that:

$$Y = \phi\theta + e \tag{2}$$

Now (2) can be substituted into (1) granting us:

$$\hat{\theta} = \theta + \underbrace{(\phi^\top \phi)^{-1} \phi^\top e}_{(4)} \tag{3}$$

We can divide (4) like this:

$$\underbrace{(\phi^\top \phi)^{-1}}_{(5)} \underbrace{\phi^\top e}_{(6)}$$

We can use Lemma B.3 on (6):

$$X_N = \frac{1}{\sqrt{N}} \sum_{t=1}^{N} \phi^\top(t) e(t) \tag{7}$$

Now as the Lemma suggests we can calculate $\lim_{N\to\infty} EX_N X_N^\top$.

$$P = \lim_{N\to\infty} EX_N X_N^\top = \lim_{N\to\infty} E\left( \frac{1}{\sqrt{N}} \sum_{t=1}^{N} \phi^\top(t) e(t) \left( \frac{1}{\sqrt{N}} \sum_{t=1}^{N} \phi^\top(t) e(t) \right)^\top \right) \tag{8}$$

Which can be rewritten as:

$$P = \lim_{N\to\infty} E\left( \frac{1}{N} \sum_{t=1}^{N} \sum_{k=1}^{N} \phi^\top(t) e(t) e^\top(k)\ \phi(k) \right) \tag{9}$$

Given that $e(t)$ and $e(k)$ are independent when $t \neq k$ the expected value of $e(t)e^\top(k)$ is zero when $t \neq k$ and $\sigma^2$ when $t = k$. This simplifies (10) to:

$$P = \lim_{N\to\infty} E\left( \frac{1}{N} \sum_{t=1}^{N} \phi^\top(t) \phi(t) \right) \sigma^2 \tag{10}$$

With Lemma B.1 (10) can be rewritten as:

$$P = \lim_{N\to\infty} E\left( \frac{1}{N} \sum_{t=1}^{N} \phi^\top(t) \phi(t) \right) \sigma^2 = E\left(E\left(\phi\phi^\top\right)\right) \sigma^2 = E\left(\phi\phi^\top\right) \sigma^2 \tag{11}$$

With the assumption that $\phi\phi^\top$ is is full rank we can show that the covariance matrix $P = E\left(\left(\phi\phi^\top\right)^{-1}\right)\sigma^2$ exists and it is therefore proven by Lemma B.3 that (6) is asymptotically gaussianally distributed. This entails that (4) is also asymptotically gaussianally distributed according to Slutsky's theorem. Slutsky's theorem also entails that since (4) is asymptotically gaussianally distributed (3) is as well. It is thereby proven that $\hat{\theta}$ is asymptotically gaussianally distributed.