

# MATH 545: INTRODUCTION TO TIME SERIES ANALYSIS

## WINTER 2025

### THE BACKSHIFT OPERATOR, CONVERGENCE, AND LINEAR FILTERS

**Backshift Operator:** The backshift operator,  $B$ , operates on the (infinite dimensional) space of finite variance discrete time series processes,  $\mathcal{X}$  say, where

$$\mathcal{X} \equiv \{X_t, t = 0, \pm 1, \pm 2, \dots, \text{Var}[X_t] < \infty\}$$

such that for  $\{X_t\} \in \mathcal{X}$ ,  $B(\{X_t\}) \equiv \{X_t^*\}$ , where, for each  $t$ ,  $X_t^* \equiv X_{t-1}$ . More informally, we write

$$BX_t = X_{t-1}.$$

$B$  is a linear operator, in that for  $\{X_t\}, \{Y_t\} \in \mathcal{X}$ , and  $a, b \in \mathbb{R}$ ,

$$B(aX_t + bY_t) = aBX_t + bBY_t = aX_{t-1} + bY_{t-1}.$$

and also has the usual algebraic properties, specifically

$$B^d X_t = B(B^{d-1} X_t) = \dots = B(B(B \dots (BX_t) \dots)) = X_{t-d}.$$

Because of these properties, we can define polynomial combinations of backshift operators, say

$$\psi_0 + \psi_1 B + \psi_2 B^2 + \dots + \psi_j B^j + \dots$$

**Linear Filters:** For arbitrary  $\{X_t\} \in \mathcal{X}$ , consider the time-homogeneous *linear filter* with real coefficients,  $\Psi(B)$ , defined by the equivalent expressions

$$\Psi(B)X_t = \left( \sum_{j=-\infty}^{\infty} \psi_j B^j \right) X_t = \sum_{j=-\infty}^{\infty} \psi_j X_{t-j}. \quad (1)$$

We first establish that for each  $t$ ,  $\Psi(B)X_t$  is finite with probability 1 (that is, this quantity is a finite random variable), and that the quantities in (1) can be considered as the almost sure limit of the (obviously well-defined) partial sums

$$\sum_{j=-k}^k \psi_j X_{t-j}. \quad (2)$$

as  $k \rightarrow \infty$ . Due to non-negativity of its terms, the sequence of partial sums indexed by  $k$

$$\sum_{j=-k}^k |\psi_j X_{t-j}| = \sum_{j=-k}^k |\psi_j| |X_{t-j}|$$

form a monotonic (non-decreasing) sequence of random variables. Therefore, for the corresponding sequence of expectations,

$$\mathbb{E} \left[ \sum_{j=-k}^k |\psi_j| |X_{t-j}| \right] \rightarrow \mathbb{E} \left[ \sum_{j=-\infty}^{\infty} |\psi_j| |X_{t-j}| \right]$$

as  $k \rightarrow \infty$ . We have

$$\mathbb{E} \left[ \left| \Psi(B)X_t \right| \right] = \mathbb{E} \left[ \left| \sum_{j=-\infty}^{\infty} \psi_j X_{t-j} \right| \right] \leq \mathbb{E} \left[ \sum_{j=-\infty}^{\infty} |\psi_j X_{t-j}| \right] \leq \sum_{j=-\infty}^{\infty} |\psi_j| \mathbb{E} [|X_{t-j}|]$$

Thus, if we specify that  $\mathbb{E} [|X_t|] < \infty$  for all  $t$ , the right hand side of the inequality is finite provided

$$\sum_{j=-\infty}^{\infty} |\psi_j| < \infty. \quad (3)$$

Note that if (3) holds, then

$$\sum_{j=-\infty}^{\infty} \psi_j \leq \sum_{j=-\infty}^{\infty} |\psi_j| < \infty$$

and also

$$\sum_{j=-\infty}^{\infty} \psi_j^2 = \sum_{j=-\infty}^{\infty} |\psi_j|^2 \leq \left( \sum_{j=-\infty}^{\infty} |\psi_j| \right)^2 < \infty. \quad (4)$$

Recall Markov's inequality: for non-negative random variable  $X$ , and constant  $c > 0$ , we have that  $P(X \geq c) \leq \mathbb{E}[X]/c$ . Thus for all  $t$

$$P(|\Psi(B)X_t| \geq c) \leq \frac{\mathbb{E}[|\Psi(B)X_t|]}{c}.$$

but as shown above, under (3),  $\mathbb{E}[|\Psi(B)X_t|] < \infty$ . Hence

$$\lim_{c \rightarrow \infty} P(|\Psi(B)X_t| \geq c) = 0$$

and so  $\Psi(B)X_t$  is finite, and the infinite sum in (1) is convergent, with probability 1.

**Mean-square convergence:** Consider the partial sums in (2). Let  $l > k > 0$ , and consider

$$\left( \sum_{j=-l}^l \psi_j X_{t-j} - \sum_{j=-k}^k \psi_j X_{t-j} \right)^2 = \left( \sum_{k < |j| \leq l} \psi_j X_{t-j} \right)^2 \quad (5)$$

Taking expectations, we have that

$$\mathbb{E} \left[ \left( \sum_{k < |j| \leq l} \psi_j X_{t-j} \right)^2 \right] = \mathbb{E} \left[ \left( \sum_{j_1} \psi_{j_1} X_{t-j_1} \right) \left( \sum_{j_2} \psi_{j_2} X_{t-j_2} \right) \right] = \sum_{j_1} \sum_{j_2} \psi_{j_1} \psi_{j_2} \mathbb{E}[X_{t-j_1} X_{t-j_2}]. \quad (6)$$

If we now also assume that  $\mathbb{E}[X_t^2] < M < \infty$  for all  $t$ , then it is also the case that  $\mathbb{E}[X_t X_s] < \infty$  for all  $t$  and  $s$  by the Cauchy-Schwarz inequality:

$$\mathbb{E}[X_t X_s] \leq \sqrt{\mathbb{E}[X_t^2] \mathbb{E}[X_s^2]}$$

Thus the expectations in (6) are also finite.

Then, with the summations over indices  $j$  such that  $k < |j| \leq l$ , from (6)

$$\begin{aligned} \mathbb{E} \left[ \left( \sum_j \psi_j X_{t-j} \right)^2 \right] &\leq M \sum_{j_1} \sum_{j_2} \psi_{j_1} \psi_{j_2} \\ &\leq M \sum_{j_1} \sum_{j_2} |\psi_{j_1} \psi_{j_2}| \\ &\leq M \left( \sum_j |\psi_j| \right)^2 \rightarrow 0 \quad k, l \rightarrow \infty \end{aligned}$$

and so the expectation of the square of the differences between the partial sums in (5) converges to zero in the limit. This is sufficient to show that the limit in (1) exists in the mean-square sense, as the partial sums form a Cauchy sequence under the mean-square discrepancy.

**Linear filters of white noise processes:** Suppose that  $\{Z_t\} \sim \text{WN}(0, \sigma_Z^2)$ . Then the linear process  $\{X_t\}$  defined in terms of  $\{Z_t\}$

$$X_t \stackrel{\text{m.s.}}{=} \sum_{j=-\infty}^{\infty} \psi_j Z_{t-j}$$

is well-defined under the assumption (3). As above in the general case

$$\mathbb{E}[|X_t|] = \mathbb{E} \left[ \left| \sum_{j=-\infty}^{\infty} \psi_j Z_{t-j} \right| \right] \leq \sum_{j=-\infty}^{\infty} |\psi_j| \mathbb{E}[|Z_{t-j}|] < \infty$$

and as  $\mathbb{E}[X_t] \leq \mathbb{E}[|X_t|]$ ,  $\mathbb{E}[X_t]$  is finite, and hence  $\mathbb{E}[X_t] = 0$  as  $\mathbb{E}[Z_t] = 0$  for all  $t$ . Secondly, due to the uncorrelatedness of the elements of the white noise process,

$$\text{Var}[X_t] = \mathbb{E}[X_t^2] = \sum_{j=-\infty}^{\infty} \psi_j^2 \mathbb{E}[Z_{t-j}^2] = \sigma_Z^2 \sum_{j=-\infty}^{\infty} \psi_j^2 < \infty.$$

Note also that

$$X_t - \sum_{j=-k}^k \psi_j Z_{t-j} = \sum_{j=-\infty}^{-k-1} \psi_j Z_{t-j} + \sum_{j=k+1}^{\infty} \psi_j Z_{t-j}$$

and hence we have that

$$\mathbb{E} \left[ \left( X_t - \sum_{j=-k}^k \psi_j Z_{t-j} \right)^2 \right] = \sigma_Z^2 \left[ \sum_{j=-\infty}^{-k-1} \psi_j^2 + \sum_{j=k+1}^{\infty} \psi_j^2 \right] \rightarrow 0$$

as  $k \rightarrow \infty$  as the right-hand side is convergent.

This confirms that in this case, under the assumption (3), the sum in (1) converges almost surely and in mean-square to a well-defined random variable  $X_t$ , with

$$\mathbb{E}[X_t] = 0 \quad \text{Var}[X_t] = \sigma_Z^2 \sum_{j=-\infty}^{\infty} \psi_j^2$$

**Generating function:** The properties of the filter may be understood through the (polynomial) *generating function* for the coefficients,  $\Psi(z)$ , which takes the form

$$\Psi(z) = \sum_{j=-\infty}^{\infty} \psi_j z^j \quad z \in \mathbb{C}.$$

For example, if  $\Psi(1)$  is finite, then

$$\sum_{j=-\infty}^{\infty} \psi_j < \infty.$$

If  $\Psi(z)$  is a polynomial of finite order including possibly terms with negative exponent, say

$$\begin{aligned} \Psi(z) &= \sum_{j=-d}^{-d+p} \psi_j z^j \\ &= \psi_{-d} z^{-d} + \psi_{-d+1} z^{-d+1} + \cdots + \psi_{-1} z^{-1} + \psi_0 + \psi_1 z + \cdots + \psi_{-d+(p-1)} z^{(p-1-d)} + \psi_{-d+p} z^{p-d} \end{aligned}$$

then we have the factorization

$$\Psi(z) = z^{-d} \prod_{j=1}^p (1 - \lambda_j z)$$

**Example:** Consider the causal AR( $p$ ) process

$$(1 - \phi_1 B - \phi_2 B^2 - \cdots - \phi_p B^p) X_t = \Phi(B) X_t = Z_t$$

defined wrt  $\{Z_t\} \sim \text{WN}(0, \sigma_Z^2)$ . The corresponding generating function, denoted  $\Phi(z)$  in this case, has factorization

$$\Phi(z) = \sum_{j=0}^p \phi_j z^j = \prod_{j=1}^p (1 - \xi_j z)$$

The linear process representation of  $\{X_t\}$ , defined via filter  $\Psi(B)$  say, can be computed using the identity

$$\Psi(B) \Phi(B) X_t = \Psi(B) Z_t$$

with the requirement that

$$\Psi(B) \Phi(B) X_t \equiv X_t.$$

In terms of the generating functions, this means  $\Psi(z) \Phi(z) = 1$  and hence we must have

$$\Psi(z) = \prod_{j=1}^p (1 - \xi_j z)^{-1}$$

Recall that for the process to be well-defined, we need that this generating function is finite when evaluated at  $|z| = 1$ ; this is equivalent to requiring that  $(1 - \xi_j z) \neq 0$ , and hence that  $\xi_j \neq 1$ , for each  $j$ .