应用时间序列分析
课程介绍

主讲老师：席瑞斌

本课件基于李东风老师课件修改

2017年秋季学期

非决定性平稳序列—介绍

- 对平稳序列，考虑用所有的历史$\{X_t, t \le n\}$对X_{n+1}进行最佳线性预测.
- 当预测误差是零时，X_{n+1}的信息完全含在历史资料中. 这样的平稳序列被称为决定性的. 2.3.5小节中Γ_{n+1}不满秩造成X_t可以被$X_{t-1}, \ldots, X_{t-n}$完全线性预测，是决定性平稳列的特例。
- 最小序列: 用$\{X_s : s \ne t\}$预报X_t误差不为零。决定性序列不是最小序列。
- 实际问题中，决定性平稳序列描述事物的发展没有新的信息出现.

非决定性平稳序列—介绍

- 对平稳序列，考虑用所有的历史$\{X_t, \ t \le n\}$对X_{n+1}进行最佳线性预测.
- 当预测误差是零时，X_{n+1}的信息完全含在历史资料中. 这样的平稳序列被称为决定性的. 2.3.5小节中Γ_{n+1}不满秩造成X_t可以被$X_{t-1}, \ldots, X_{t-n}$完全线性预测，是决定性平稳列的特例。
- 最小序列：用$\{X_s : s \neq t\}$预报X_t误差不为零。决定性序列不是最小序列。
- 实际问题中，决定性平稳序列描述事物的发展没有新的信息出现.

非决定性平稳序列—介绍

- 对平稳序列，考虑用所有的历史$\{X_t, \ t \leq n\}$对X_{n+1}进行最佳线性预测.
- 当预测误差是零时，X_{n+1}的信息完全含在历史资料中. 这样的平稳序列被称为决定性的. 2.3.5小节中Γ_{n+1}不满秩造成X_t可以被$X_{t-1}, \ldots, X_{t-n}$完全线性预测，是决定性平稳列的特例。
- 最小序列: 用$\{X_s : s \neq t\}$预报X_t误差不为零。决定性序列不是最小序列。
- 实际问题中，决定性平稳序列描述事物的发展没有新的信息出现.

非决定性平稳序列—介绍

- 对平稳序列，考虑用所有的历史$\{X_t, \ t \le n\}$对X_{n+1}进行最佳线性预测.
- 当预测误差是零时，X_{n+1}的信息完全含在历史资料中. 这样的平稳序列被称为决定性的. 2.3.5小节中Γ_{n+1}不满秩造成X_t可以被$X_{t-1}, \ldots, X_{t-n}$完全线性预测，是决定性平稳列的特例。
- 最小序列: 用$\{X_s : s \neq t\}$预报X_t误差不为零。决定性序列不是最小序列。
- 实际问题中，决定性平稳序列描述事物的发展没有新的信息出现.

- 如果用$\{X_t, t \leq n\}$预测X_{n+1}的误差不是零，说明X_{n+1}的信息不能由历史资料的线性组合及其极限完全确定，我们称这种时间序列是非决定性的.
- 非决定性平稳序列描述事物的发展总伴随新的信息出现.
- 最小序列一定是非决定性的。
- 平稳序列的Wold定理表示告诉我们，非决定性平稳序列总是可以分解成白噪声的单边滑动和加上一个决定性平稳序列.
- 从应用的角度讲，非决定性平稳序列总是白噪声的单边滑动和加上一个离散谱序列.

- 如果用$\{X_t, t \leq n\}$预测X_{n+1}的误差不是零，说明X_{n+1}的信息不能由历史资料的线性组合及其极限完全确定，我们称这种时间序列是非决定性的.
- 非决定性平稳序列描述事物的发展总伴随新的信息出现.
- 最小序列一定是非决定性的。
- 平稳序列的Wold定理表示告诉我们，非决定性平稳序列总是可以分解成白噪声的单边滑动和加上一个决定性平稳序列.
- 从应用的角度讲，非决定性平稳序列总是白噪声的单边滑动和加上一个离散谱序列.

- 如果用$\{X_t, t \leq n\}$预测X_{n+1}的误差不是零，说明X_{n+1}的信息不能由历史资料的线性组合及其极限完全确定，我们称这种时间序列是非决定性的.
- 非决定性平稳序列描述事物的发展总伴随新的信息出现.
- 最小序列一定是非决定性的。
- 平稳序列的Wold定理表示告诉我们，非决定性平稳序列总是可以分解成白噪声的单边滑动和加上一个决定性平稳序列.
- 从应用的角度讲，非决定性平稳序列总是白噪声的单边滑动和加上一个离散谱序列.

- 如果用$\{X_t, t \le n\}$预测X_{n+1}的误差不是零，说明X_{n+1}的信息不能由历史资料的线性组合及其极限完全确定，我们称这种时间序列是非决定性的.
- 非决定性平稳序列描述事物的发展总伴随新的信息出现.
- 最小序列一定是非决定性的。
- 平稳序列的Wold定理表示告诉我们，非决定性平稳序列总是可以分解成白噪声的单边滑动和加上一个决定性平稳序列.
- 从应用的角度讲，非决定性平稳序列总是白噪声的单边滑动和加上一个离散谱序列.

- 如果用$\{X_t, t \le n\}$预测X_{n+1}的误差不是零，说明X_{n+1}的信息不能由历史资料的线性组合及其极限完全确定，我们称这种时间序列是非决定性的.
- 非决定性平稳序列描述事物的发展总伴随新的信息出现.
- 最小序列一定是非决定性的。
- 平稳序列的Wold定理表示告诉我们，非决定性平稳序列总是可以分解成白噪声的单边滑动和加上一个决定性平稳序列.
- 从应用的角度讲，非决定性平稳序列总是白噪声的单边滑动和加上一个离散谱序列.

最佳线性预测均方误差的极限

- 设$\{X_n : n \in \mathbb{Z}\}$是零均值平稳序列. 记

$$\mathbf{X}_{n,m} = (X_n, X_{n-1}, \cdots, X_{n-m+1})^T,$$

这里n表示向量的第一个脚标, m表示向量的维数.

- 定义

$$\hat{X}_{n+1,m} = L(X_{n+1}|\mathbf{X}_{n,m}).$$

- 从最佳线性预测的性质8知道$\sigma_{1,m}^2 = E(X_{n+1} - \hat{X}_{n+1,m})^2$ 是m的单调减函数,于是定义

$$\sigma_1^2 \triangleq \lim_{m\to\infty} \sigma_{1,m}^2 < \infty.$$

最佳线性预测均方误差的极限

- 设$\{X_n : n \in \mathbb{Z}\}$是零均值平稳序列. 记

$$\mathbf{X}_{n,m} = (X_n, X_{n-1}, \cdots, X_{n-m+1})^T,$$

这里n表示向量的第一个脚标, m表示向量的维数.

- 定义

$$\hat{X}_{n+1,m} = L(X_{n+1}|\mathbf{X}_{n,m}).$$

- 从最佳线性预测的性质8知道$\sigma_{1,m}^2 = E(X_{n+1} - \hat{X}_{n+1,m})^2$ 是m的单调减函数,于是定义

$$\sigma_1^2 \triangleq \lim_{m\to\infty} \sigma_{1,m}^2 < \infty.$$

最佳线性预测均方误差的极限

- 设$\{X_n : n \in \mathbb{Z}\}$是零均值平稳序列. 记

$$\mathbf{X}_{n,m} = (X_n, X_{n-1}, \cdots, X_{n-m+1})^T,$$

这里n表示向量的第一个脚标, m表示向量的维数.

- 定义

$$\hat{X}_{n+1,m} = L(X_{n+1}|\mathbf{X}_{n,m}).$$

- 从最佳线性预测的性质8知道$\sigma_{1,m}^2 = E(X_{n+1} - \hat{X}_{n+1,m})^2$ 是m的单调减函数,于是定义

$$\sigma_1^2 \triangleq \lim_{m\to\infty} \sigma_{1,m}^2 < \infty.$$

最佳线性预测均方误差的极限II

- **定理2.1** $\sigma_1^2 \triangleq \lim_{m\to\infty} \sigma_{1,m}^2$与$n$无关.
- **证明** 设$\mathbf{a} = (a_1, a_2, \ldots, a_m)^T$ 是预测方程(1.5)的解, 则$\mathbf{a}$和n无关. 由于

$$Y_n \triangleq X_{n+1} - \sum_{j=1}^{m} a_j X_{n+1-j} = X_{n+1} - \hat{X}_{n+1,m}, \quad n \in \mathbb{Z}, \tag{2.1}$$

是平稳序列, 所以$\sigma_{1,m}^2 = EY_n^2 = EY_0^2$与$n$无关. 最后$\sigma_1^2 = \lim_{m\to\infty} \sigma_{1,m}^2$与$n$无关. □

最佳线性预测均方误差的极限II

- **定理2.1** $\sigma_1^2 \triangleq \lim_{m\to\infty}\sigma_{1,m}^2$与$n$无关.
- **证明** 设$\mathbf{a}=(a_1,a_2,\ldots,a_m)^T$ 是预测方程(1.5)的解, 则$\mathbf{a}$和n无关. 由于

$$Y_n \triangleq X_{n+1}-\sum_{j=1}^{m} a_j X_{n+1-j} = X_{n+1}-\hat{X}_{n+1,m}, \quad n\in\mathbb{Z}, \tag{2.1}$$

是平稳序列, 所以$\sigma_{1,m}^2 = EY_n^2 = EY_0^2$与$n$无关. 最后$\sigma_1^2=\lim_{m\to\infty}\sigma_{1,m}^2$与$n$无关. □

决定性与非决定性

- 对充分大的m, $L(X_{n+1}|\mathbf{X}_{n,m})$表示用充分多的历史对未来$X_{n+1}$进行预测.
- $\sigma^2_{1,m}$表示的是预测的均方误差.
- 当$m \to \infty$时, $\sigma^2_{1,m} \to 0$ 说明X_{n+1} 可以由所有历史$X_n, X_{n-1}, \ldots$进行完全预测.
- 当$\sigma^2_1 > 0$ 说明X_{n+1}不可以由所有历史$X_n, X_{n-1}, \ldots$的线性组合以及极限进行完全预测.
- **定义2.1** 设$\{X_t\}$是零均值平稳序列.
  1. 如果$\sigma^2_1 = 0$,称$\{X_t\}$是决定性平稳序列;
  2. 如果$\sigma^2_1 > 0$, 称$\{X_t\}$是非决定性平稳序列, 并且称$\sigma^2_1 = \lim_{m\to\infty} \sigma^2_{1,m}$ 为$\{X_t\}$的一步(线性)预测的均方误差。

决定性与非决定性

- 对充分大的m, $L(X_{n+1}|\mathbf{X}_{n,m})$表示用充分多的历史对未来$X_{n+1}$进行预测.
- $\sigma^2_{1,m}$表示的是预测的均方误差.
- 当$m \to \infty$时, $\sigma^2_{1,m} \to 0$ 说明X_{n+1} 可以由所有历史$X_n, X_{n-1}, \ldots$进行完全预测.
- 当$\sigma_1^2 > 0$ 说明X_{n+1}不可以由所有历史$X_n, X_{n-1}, \ldots$的线性组合以及极限进行完全预测.
- **定义2.1** 设$\{X_t\}$是零均值平稳序列.
  1. 如果$\sigma_1^2 = 0$,称$\{X_t\}$是决定性平稳序列;
  2. 如果$\sigma_1^2 > 0$, 称$\{X_t\}$是非决定性平稳序列, 并且称$\sigma_1^2 = \lim_{m\to\infty} \sigma^2_{1,m}$ 为$\{X_t\}$的一步(线性)预测的均方误差。

决定性与非决定性

- 对充分大的m，$L(X_{n+1}|\mathbf{X}_{n,m})$表示用充分多的历史对未来$X_{n+1}$进行预测.
- $\sigma_{1,m}^2$表示的是预测的均方误差.
- 当$m\to\infty$时，$\sigma_{1,m}^2\to 0$ 说明X_{n+1} 可以由所有历史$X_n, X_{n-1},\ldots$进行完全预测.
- 当$\sigma_1^2>0$ 说明X_{n+1}不可以由所有历史$X_n, X_{n-1},\ldots$的线性组合以及极限进行完全预测.
- **定义2.1** 设$\{X_t\}$是零均值平稳序列.
  1. 如果$\sigma_1^2=0$,称$\{X_t\}$是决定性平稳序列;
  2. 如果$\sigma_1^2>0$, 称$\{X_t\}$是非决定性平稳序列, 并且称$\sigma_1^2=\lim_{m\to\infty}\sigma_{1,m}^2$ 为$\{X_t\}$的一步(线性)预测的均方误差。

决定性与非决定性

- 对充分大的m, $L(X_{n+1}|\mathbf{X}_{n,m})$表示用充分多的历史对未来$X_{n+1}$进行预测.
- $\sigma^2_{1,m}$表示的是预测的均方误差.
- 当$m \to \infty$时, $\sigma^2_{1,m} \to 0$ 说明X_{n+1} 可以由所有历史$X_n, X_{n-1}, \ldots$进行完全预测.
- 当$\sigma_1^2 > 0$ 说明X_{n+1}不可以由所有历史$X_n, X_{n-1}, \ldots$的线性组合以及极限进行完全预测.
- **定义2.1** 设$\{X_t\}$是零均值平稳序列.
  1. 如果$\sigma_1^2 = 0$,称$\{X_t\}$是决定性平稳序列;
  2. 如果$\sigma_1^2 > 0$, 称$\{X_t\}$是非决定性平稳序列, 并且称$\sigma_1^2 = \lim_{m\to\infty} \sigma^2_{1,m}$ 为$\{X_t\}$的一步(线性)预测的均方误差。

决定性与非决定性

- 对充分大的m, $L(X_{n+1}|\mathbf{X}_{n,m})$表示用充分多的历史对未来$X_{n+1}$进行预测.
- $\sigma^2_{1,m}$表示的是预测的均方误差.
- 当$m \to \infty$时, $\sigma^2_{1,m} \to 0$ 说明X_{n+1} 可以由所有历史$X_n, X_{n-1}, \ldots$进行完全预测.
- 当$\sigma^2_1 > 0$ 说明X_{n+1}不可以由所有历史$X_n, X_{n-1}, \ldots$的线性组合以及极限进行完全预测.
- **定义2.1** 设$\{X_t\}$是零均值平稳序列.
  1. 如果$\sigma^2_1 = 0$,称$\{X_t\}$是决定性平稳序列;
  2. 如果$\sigma^2_1 > 0$, 称$\{X_t\}$是非决定性平稳序列, 并且称$\sigma^2_1 = \lim_{m\to\infty} \sigma^2_{1,m}$ 为$\{X_t\}$的一步(线性)预测的均方误差。

决定性与非决定性

- 对充分大的m, $L(X_{n+1}|\mathbf{X}_{n,m})$表示用充分多的历史对未来$X_{n+1}$进行预测.
- $\sigma^2_{1,m}$表示的是预测的均方误差.
- 当$m \to \infty$时, $\sigma^2_{1,m} \to 0$ 说明X_{n+1} 可以由所有历史$X_n, X_{n-1}, \ldots$进行完全预测.
- 当$\sigma_1^2 > 0$ 说明X_{n+1}不可以由所有历史$X_n, X_{n-1}, \ldots$的线性组合以及极限进行完全预测.
- **定义2.1** 设$\{X_t\}$是零均值平稳序列.
  1. 如果$\sigma_1^2 = 0$,称$\{X_t\}$是**决定性平稳序列**;
  2. 如果$\sigma_1^2 > 0$, 称$\{X_t\}$是**非决定性平稳序列**, 并且称$\sigma_1^2 = \lim_{m\to\infty} \sigma^2_{1,m}$ 为$\{X_t\}$的**一步(线性)预测的均方误差**。

决定性与非决定性

- 对充分大的m，$L(X_{n+1}|\mathbf{X}_{n,m})$表示用充分多的历史对未来$X_{n+1}$进行预测.
- $\sigma^2_{1,m}$表示的是预测的均方误差.
- 当$m\to\infty$时，$\sigma^2_{1,m}\to 0$ 说明X_{n+1} 可以由所有历史$X_n, X_{n-1}, \ldots$进行完全预测.
- 当$\sigma_1^2 > 0$ 说明X_{n+1}不可以由所有历史$X_n, X_{n-1}, \ldots$的线性组合以及极限进行完全预测.
- **定义2.1** 设$\{X_t\}$是零均值平稳序列.
  1. 如果$\sigma_1^2 = 0$,称$\{X_t\}$是**决定性平稳序列**;
  2. 如果$\sigma_1^2 > 0$，称$\{X_t\}$是**非决定性平稳序列**，并且称$\sigma_1^2 = \lim_{m\to\infty}\sigma^2_{1,m}$ 为$\{X_t\}$的**一步(线性)预测的均方误差**。

非零均值情况

- 对于平稳序列$\{X_t\}$，如果$EX_t=\mu$，引入$\{Z_t\}=\{X_t-\mu\}$和m维向量$\boldsymbol{\mu}_m=(\mu,\ldots,\mu)^T$.
- 按照最佳线性预测的定义1.2，

$$\begin{aligned}\hat{X}_{n+1,m}=&\mu+L(X_{n+1}-\mu|\mathbf{X}_{n,m}-\boldsymbol{\mu}_m)\\=&\mu+L(Z_{n+1}|\mathbf{Z}_{n,m})=\mu+\hat{Z}_{n+1,m}.\end{aligned}$$

- 于是

$$E(Z_{n+1}-\hat{Z}_{n+1,m})^2=E(X_{n+1}-\hat{X}_{n+1,m})^2.$$

即对X预报的均方误差等于对中心化得到的Z预报的均方误差。

- 因而，当且仅当$\{X_t-\mu\}$是决定性平稳序列时，称$\{X_t\}$是决定性平稳序列．于是以后只需要讨论零均值的平稳序列.

非零均值情况

- 对于平稳序列$\{X_t\}$，如果$EX_t = \mu$，引入$\{Z_t\} = \{X_t - \mu\}$ 和m维向量$\boldsymbol{\mu}_m = (\mu, \ldots, \mu)^T$.
- 按照最佳线性预测的定义1.2,

$$\begin{aligned}\hat{X}_{n+1,m} =& \mu + L(X_{n+1} - \mu|\mathbf{X}_{n,m} - \boldsymbol{\mu}_m) \\ =& \mu + L(Z_{n+1}|\mathbf{Z}_{n,m}) = \mu + \hat{Z}_{n+1,m}.\end{aligned}$$

- 于是

$$E(Z_{n+1} - \hat{Z}_{n+1,m})^2 = E(X_{n+1} - \hat{X}_{n+1,m})^2.$$

即对X预报的均方误差等于对中心化得到的Z预报的均方误差。

- 因而，当且仅当$\{X_t - \mu\}$是决定性平稳序列时，称$\{X_t\}$是决定性平稳序列. 于是以后只需要讨论零均值的平稳序列.

非零均值情况

- 对于平稳序列$\{X_t\}$，如果$EX_t = \mu$，引入$\{Z_t\} = \{X_t - \mu\}$ 和m维向量$\boldsymbol{\mu}_m = (\mu, \ldots, \mu)^T$.
- 按照最佳线性预测的定义1.2，

$$
\begin{aligned}
\hat{X}_{n+1,m} =& \mu + L(X_{n+1} - \mu | \mathbf{X}_{n,m} - \boldsymbol{\mu}_m) \\
=& \mu + L(Z_{n+1} | \mathbf{Z}_{n,m}) = \mu + \hat{Z}_{n+1,m}.
\end{aligned}
$$

- 于是

$$
E(Z_{n+1} - \hat{Z}_{n+1,m})^2 = E(X_{n+1} - \hat{X}_{n+1,m})^2.
$$

即对X预报的均方误差等于对中心化得到的Z预报的均方误差。

- 因而，当且仅当$\{X_t - \mu\}$是决定性平稳序列时，称$\{X_t\}$是决定性平稳序列. 于是以后只需要讨论零均值的平稳序列.

非零均值情况

- 对于平稳序列$\{X_t\}$，如果$EX_t = \mu$，引入$\{Z_t\} = \{X_t - \mu\}$ 和m维向量$\boldsymbol{\mu}_m = (\mu, \ldots, \mu)^T$.
- 按照最佳线性预测的定义1.2，

$$\begin{aligned}\hat{X}_{n+1,m} =& \mu + L(X_{n+1} - \mu|\mathbf{X}_{n,m} - \boldsymbol{\mu}_m) \\ =& \mu + L(Z_{n+1}|\mathbf{Z}_{n,m}) = \mu + \hat{Z}_{n+1,m}.\end{aligned}$$

- 于是

$$E(Z_{n+1} - \hat{Z}_{n+1,m})^2 = E(X_{n+1} - \hat{X}_{n+1,m})^2.$$

  即对X预报的均方误差等于对中心化得到的Z预报的均方误差。
- 因而，当且仅当$\{X_t - \mu\}$是决定性平稳序列时，称$\{X_t\}$是决定性平稳序列．于是以后只需要讨论零均值的平稳序列.

例2.1 可完全线性预测

- 设平稳列$\{X_t\}$的$n+1$阶自协方差阵Γ_{n+1}退化，$|\Gamma_n|>0$。
- 则$X_1, X_2, \cdots, X_{n+1}$线性相关，
- 所以X_{n+1}可以由$X_n, X_{n-1}, \ldots, X_1$线性表示.
- 于是，$L(X_{n+1}|X_n, \ldots, X_1) = X_{n+1}$.
- 当$m \geq n$时，

$$L(X_{n+1}|X_n, \ldots, X_{n-m+1}) = X_{n+1},$$

- 即有$\sigma^2_{1,m} = 0$, $\{X_t\}$是决定性平稳列。
- 最简单的决定性平稳列是$X_t \equiv \xi$, ξ为随机变量。

例2.1 可完全线性预测

- 设平稳列$\{X_t\}$的$n+1$阶自协方差阵Γ_{n+1}退化，$|\Gamma_n|>0$。
- 则$X_1, X_2, \cdots, X_{n+1}$线性相关，
- 所以X_{n+1}可以由$X_n, X_{n-1}, \ldots, X_1$线性表示.
- 于是，$L(X_{n+1}|X_n, \ldots, X_1) = X_{n+1}$.
- 当$m \geq n$时，

$$L(X_{n+1}|X_n, \ldots, X_{n-m+1}) = X_{n+1},$$

- 即有$\sigma^2_{1,m} = 0$，$\{X_t\}$是决定性平稳列。
- 最简单的决定性平稳列是$X_t \equiv \xi$，ξ为随机变量。

例2.1 可完全线性预测

- 设平稳列$\{X_t\}$的$n+1$阶自协方差阵Γ_{n+1}退化，$|\Gamma_n|>0$。
- 则$X_1,X_2,\cdots,X_{n+1}$线性相关,
- 所以X_{n+1}可以由$X_n,X_{n-1},\ldots,X_1$线性表示.
- 于是, $L(X_{n+1}|X_n,\ldots,X_1)=X_{n+1}$.
- 当$m\geq n$时,

$$L(X_{n+1}|X_n,\ldots,X_{n-m+1})=X_{n+1},$$

- 即有$\sigma^2_{1,m}=0$, $\{X_t\}$是决定性平稳列。
- 最简单的决定性平稳列是$X_t\equiv\xi$, ξ为随机变量。

例2.1 可完全线性预测

- 设平稳列$\{X_t\}$的$n+1$阶自协方差阵Γ_{n+1}退化，$|\Gamma_n|>0$。
- 则$X_1,X_2,\cdots,X_{n+1}$线性相关,
- 所以X_{n+1}可以由$X_n,X_{n-1},\ldots,X_1$线性表示.
- 于是, $L(X_{n+1}|X_n,\ldots,X_1)=X_{n+1}$.
- 当$m\geq n$时,

$$L(X_{n+1}|X_n,\ldots,X_{n-m+1})=X_{n+1},$$

- 即有$\sigma^2_{1,m}=0$, $\{X_t\}$是决定性平稳列。
- 最简单的决定性平稳列是$X_t\equiv\xi$, ξ为随机变量。

例2.1 可完全线性预测

- 设平稳列$\{X_t\}$的$n+1$阶自协方差阵Γ_{n+1}退化，$|\Gamma_n|>0$。
- 则$X_1, X_2, \cdots, X_{n+1}$线性相关,
- 所以X_{n+1}可以由$X_n, X_{n-1}, \ldots, X_1$线性表示.
- 于是, $L(X_{n+1}|X_n, \ldots, X_1) = X_{n+1}$.
- 当$m \geq n$时,

$$L(X_{n+1}|X_n, \ldots, X_{n-m+1}) = X_{n+1},$$

- 即有$\sigma^2_{1,m} = 0$, $\{X_t\}$是决定性平稳列。
- 最简单的决定性平稳列是$X_t \equiv \xi$, ξ为随机变量。

例2.1 可完全线性预测

- 设平稳列$\{X_t\}$的$n+1$阶自协方差阵Γ_{n+1}退化，$|\Gamma_n|>0$。
- 则$X_1, X_2, \cdots, X_{n+1}$线性相关,
- 所以X_{n+1}可以由$X_n, X_{n-1}, \ldots, X_1$线性表示.
- 于是, $L(X_{n+1}|X_n, \ldots, X_1) = X_{n+1}$.
- 当$m \geq n$时,

$$L(X_{n+1}|X_n, \ldots, X_{n-m+1}) = X_{n+1},$$

- 即有$\sigma^2_{1,m} = 0$, $\{X_t\}$是决定性平稳列。
- 最简单的决定性平稳列是$X_t \equiv \xi$, ξ为随机变量。

例2.1 可完全线性预测

- 设平稳列$\{X_t\}$的$n+1$阶自协方差阵Γ_{n+1}退化，$|\Gamma_n|>0$。
- 则$X_1, X_2, \cdots, X_{n+1}$线性相关,
- 所以X_{n+1}可以由$X_n, X_{n-1}, \ldots, X_1$线性表示.
- 于是, $L(X_{n+1}|X_n, \ldots, X_1) = X_{n+1}$.
- 当$m \geq n$时,

$$L(X_{n+1}|X_n, \ldots, X_{n-m+1}) = X_{n+1},$$

- 即有$\sigma^2_{1,m} = 0$, $\{X_t\}$是决定性平稳列。
- 最简单的决定性平稳列是$X_t \equiv \xi$, ξ为随机变量。

例2.2 离散谱序列

- 设零均值随机变量ξ_j, η_k $(j,k=1,2,\ldots,p)$ 两两正交，满足

$$E(\xi_j^2)=E(\eta_j^2)=\sigma_j^2,\ j=1,2,\ldots \qquad (2.2)$$

- 对确定的j，定义简单离散谱序列

$$Z_j(t)=\xi_j\cos(t\lambda_j)+\eta_j\sin(t\lambda_j),\quad t\in\mathbb{Z}. \qquad (2.3)$$

- 可以证明$\{Z_j(t)\}$是平稳序列。
- 事实上，易见$\mathrm{E}Z_j(t)\equiv 0$。而

$$\begin{aligned}\mathrm{E}[Z_j(t)Z_j(s)]=&\mathrm{E}\xi_j^2\cos(\lambda_j t)\cos(\lambda_j s)+\mathrm{E}\eta_j^2\sin(\lambda_j t)\sin(\lambda_j s)\\=&\sigma_j^2\cos((t-s)\lambda_j)\end{aligned}$$

只依赖于$t-s$。

例2.2 离散谱序列

- 设零均值随机变量ξ_j, η_k $(j,k=1,2,\dots,p)$ 两两正交，满足

$$E(\xi_j^2)=E(\eta_j^2)=\sigma_j^2,\ j=1,2,\dots \qquad (2.2)$$

- 对确定的j，定义简单离散谱序列

$$Z_j(t)=\xi_j\cos(t\lambda_j)+\eta_j\sin(t\lambda_j),\quad t\in\mathbb{Z}. \qquad (2.3)$$

- 可以证明$\{Z_j(t)\}$是平稳序列。
- 事实上，易见$\mathrm{E}Z_j(t)=0$。而

$$\begin{aligned}\mathrm{E}[Z_j(t)Z_j(s)]=&\mathrm{E}\xi_j^2\cos(\lambda_j t)\cos(\lambda_j s)+\mathrm{E}\eta_j^2\sin(\lambda_j t)\sin(\lambda_j s)\\=&\sigma_j^2\cos((t-s)\lambda_j)\end{aligned}$$

只依赖于$t-s$。

例2.2 离散谱序列

- 设零均值随机变量ξ_j, η_k $(j,k=1,2,\ldots,p)$ 两两正交，满足

$$E(\xi_j^2)=E(\eta_j^2)=\sigma_j^2,\ j=1,2,\ldots \qquad (2.2)$$

- 对确定的j, 定义简单离散谱序列

$$Z_j(t)=\xi_j\cos(t\lambda_j)+\eta_j\sin(t\lambda_j),\quad t\in\mathbb{Z}. \qquad (2.3)$$

- 可以证明$\{Z_j(t)\}$是平稳序列。
- 事实上，易见$\mathrm{E}Z_j(t)\equiv 0$。而

$$\begin{aligned}\mathrm{E}[Z_j(t)Z_j(s)]=&\mathrm{E}\xi_j^2\cos(\lambda_j t)\cos(\lambda_j s)+\mathrm{E}\eta_j^2\sin(\lambda_j t)\sin(\lambda_j s)\\=&\sigma_j^2\cos((t-s)\lambda_j)\end{aligned}$$

只依赖于$t-s$。

例2.2 离散谱序列

- 设零均值随机变量ξ_j, η_k $(j, k = 1, 2, \ldots, p)$ 两两正交，满足

$$E(\xi_j^2) = E(\eta_j^2) = \sigma_j^2,\ j = 1, 2, \ldots \tag{2.2}$$

- 对确定的j，定义简单离散谱序列

$$Z_j(t) = \xi_j \cos(t\lambda_j) + \eta_j \sin(t\lambda_j), \quad t \in \mathbb{Z}. \tag{2.3}$$

- 可以证明$\{Z_j(t)\}$是平稳序列。
- 事实上，易见$\mathrm{E}Z_j(t) \equiv 0$。而

$$\begin{aligned} \mathrm{E}[Z_j(t)Z_j(s)] =& \mathrm{E}\xi_j^2 \cos(\lambda_j t)\cos(\lambda_j s) + \mathrm{E}\eta_j^2 \sin(\lambda_j t)\sin(\lambda_j s) \\ =& \sigma_j^2 \cos((t-s)\lambda_j) \end{aligned}$$

只依赖于$t - s$。

例2.2 离散谱序列II

- $\{Z_j(t)\}$的每一次实现是周期函数，由§2.3的定理3.7 知道$\{Z_j(t)\}$的3阶自协方差矩阵是退化的，
- 因此(2.3)是可完全线性预测的，$Z_j(n)$可以被$X_j(n-1), X_j(n-2)$完全线性预测，是决定性序列。
- 事实上容易证明

$$Z_j(t) = (2\cos\lambda_j)Z_j(t-1) - Z_j(t-2)$$

- 定义离散谱序列

$$Z_t = \sum_{j=1}^{p} Z_j(t), \quad t \in \mathbb{Z}. \tag{2.4}$$

这是p个简单离散谱序列的叠加.

- 由§2.3的定理3.7知道由(2.4)定义的离散谱序列也是决定性的. Z_n可以被$Z_{n-1}, Z_{n-2}, \ldots, Z_{n-2p}$完全线性预测。

例2.2 离散谱序列II

- $\{Z_j(t)\}$的每一次实现是周期函数，由§2.3的定理3.7 知道$\{Z_j(t)\}$的3阶自协方差矩阵是退化的，
- 因此(2.3)是可完全线性预测的，$Z_j(n)$可以被$X_j(n-1), X_j(n-2)$完全线性预测，是决定性序列。
- 事实上容易证明

$$Z_j(t) = (2\cos\lambda_j)Z_j(t-1) - Z_j(t-2)$$

- 定义离散谱序列

$$Z_t = \sum_{j=1}^{p} Z_j(t), \quad t \in \mathbb{Z}. \tag{2.4}$$

这是p个简单离散谱序列的叠加.

- 由§2.3的定理3.7知道由(2.4)定义的离散谱序列也是决定性的. Z_n可以被$Z_{n-1}, Z_{n-2}, \ldots, Z_{n-2p}$完全线性预测。

例2.2 离散谱序列II

- $\{Z_j(t)\}$的每一次实现是周期函数，由§2.3的定理3.7 知道$\{Z_j(t)\}$的3阶自协方差矩阵是退化的，
- 因此(2.3)是可完全线性预测的，$Z_j(n)$可以被$X_j(n-1), X_j(n-2)$完全线性预测，是决定性序列。
- 事实上容易证明

$$Z_j(t) = (2\cos\lambda_j)Z_j(t-1) - Z_j(t-2)$$

- 定义离散谱序列

$$Z_t = \sum_{j=1}^{p} Z_j(t), \quad t \in \mathbb{Z}. \qquad (2.4)$$

这是p个简单离散谱序列的叠加.

- 由§2.3的定理3.7知道由(2.4)定义的离散谱序列也是决定性的. Z_n可以被$Z_{n-1}, Z_{n-2}, \ldots, Z_{n-2p}$完全线性预测。

例2.2 离散谱序列II

- $\{Z_j(t)\}$的每一次实现是周期函数, 由§2.3的定理3.7 知道$\{Z_j(t)\}$的3阶自协方差矩阵是退化的,
- 因此(2.3)是可完全线性预测的，$Z_j(n)$可以被$X_j(n-1), X_j(n-2)$完全线性预测，是决定性序列。
- 事实上容易证明

$$Z_j(t) = (2\cos\lambda_j)Z_j(t-1) - Z_j(t-2)$$

- 定义离散谱序列

$$Z_t = \sum_{j=1}^{p} Z_j(t), \quad t \in \mathbb{Z}. \tag{2.4}$$

这是p个简单离散谱序列的叠加.

- 由§2.3的定理3.7知道由(2.4)定义的离散谱序列也是决定性的. Z_n可以被$Z_{n-1}, Z_{n-2}, \ldots, Z_{n-2p}$完全线性预测。

例2.2 离散谱序列II

- $\{Z_j(t)\}$的每一次实现是周期函数，由§2.3的定理3.7 知道$\{Z_j(t)\}$的3阶自协方差矩阵是退化的，
- 因此(2.3)是可完全线性预测的，$Z_j(n)$可以被$X_j(n-1), X_j(n-2)$完全线性预测，是决定性序列。
- 事实上容易证明

$$Z_j(t) = (2\cos\lambda_j)Z_j(t-1) - Z_j(t-2)$$

- 定义离散谱序列

$$Z_t = \sum_{j=1}^{p} Z_j(t), \quad t \in \mathbb{Z}. \tag{2.4}$$

  这是p个简单离散谱序列的叠加.
- 由§2.3的定理3.7知道由(2.4)定义的离散谱序列也是决定性的. Z_n可以被$Z_{n-1}, Z_{n-2}, \ldots, Z_{n-2p}$完全线性预测。

纯非决定性

- 决定性与非决定性取决于一步线性预报误差是否为零。
- 对非决定性序列，用$\{X_s, s \le n\}$预报X_{n+k}的误差会随k增大而增大。
- 记

$$\sigma_{k,m}^2 = \mathrm{E}[X_{n+k} - L(X_{n+k}|X_n, X_{n-1}, \ldots, X_{n-m+1})]^2$$

则$\sigma_{k,m}$也是m的单调递减函数，与n无关。可定义

$$\sigma_k^2 = \lim_{m\to\infty} \sigma_{k,m}^2$$

- 在极限意义下可以证明$\sigma_k^2 \ge \sigma_{k-1}^2$:

$$\begin{aligned}
\sigma_k^2 &= \lim_{m\to\infty} \mathrm{E}(X_{n+k} - L(X_{n+k}|X_n, X_{n-1}, \ldots, X_{n-m}))^2 \\
&= \lim_{m\to\infty} \mathrm{E}[X_{n+k-1} - L(X_{n+k-1}|X_{n-1}, X_{n-2}, \ldots, X_{n-1-m})]^2 \\
&\ge \lim_{m\to\infty} \mathrm{E}[X_{n+k-1} - L(X_{n+k-1}|X_n, X_{n-1}, \ldots, X_{n-m-1})]^2 \\
&= \lim_{m\to\infty} \sigma_{k-1,m+1}^2 = \sigma_{k-1}^2. \qquad (2.6)
\end{aligned}$$

纯非决定性

- 决定性与非决定性取决于一步线性预报误差是否为零。
- 对非决定性序列，用$\{X_s, s \le n\}$预报X_{n+k}的误差会随k增大而增大。
- 记

$$\sigma_{k,m}^2 = \mathrm{E}[X_{n+k} - L(X_{n+k}|X_n, X_{n-1}, \ldots, X_{n-m+1})]^2$$

则$\sigma_{k,m}$也是m的单调递减函数，与n无关。可定义

$$\sigma_k^2 = \lim_{m\to\infty} \sigma_{k,m}^2$$

- 在极限意义下可以证明$\sigma_k^2 \ge \sigma_{k-1}^2$:

$$\begin{aligned}
\sigma_k^2 &= \lim_{m\to\infty} \mathrm{E}(X_{n+k} - L(X_{n+k}|X_n, X_{n-1}, \ldots, X_{n-m}))^2 \\
&= \lim_{m\to\infty} \mathrm{E}[X_{n+k-1} - L(X_{n+k-1}|X_{n-1}, X_{n-2}, \ldots, X_{n-1-m})]^2 \\
&\ge \lim_{m\to\infty} \mathrm{E}[X_{n+k-1} - L(X_{n+k-1}|X_n, X_{n-1}, \ldots, X_{n-m-1})]^2 \\
&= \lim_{m\to\infty} \sigma_{k-1,m+1}^2 = \sigma_{k-1}^2. \qquad (2.6)
\end{aligned}$$

纯非决定性

- 决定性与非决定性取决于一步线性预报误差是否为零。
- 对非决定性序列，用$\{X_s, s \le n\}$预报X_{n+k}的误差会随k增大而增大。
- 记

$$\sigma_{k,m}^2 = \mathrm{E}[X_{n+k} - L(X_{n+k}|X_n, X_{n-1}, \ldots, X_{n-m+1})]^2$$

则$\sigma_{k,m}$也是m的单调递减函数，与n无关。可定义

$$\sigma_k^2 = \lim_{m\to\infty} \sigma_{k,m}^2$$

- 在极限意义下可以证明$\sigma_k^2 \ge \sigma_{k-1}^2$:

$$\begin{aligned}
\sigma_k^2 &= \lim_{m\to\infty} \mathrm{E}(X_{n+k} - L(X_{n+k}|X_n, X_{n-1}, \ldots, X_{n-m}))^2 \\
&= \lim_{m\to\infty} \mathrm{E}[X_{n+k-1} - L(X_{n+k-1}|X_{n-1}, X_{n-2}, \ldots, X_{n-1-m})]^2 \\
&\ge \lim_{m\to\infty} \mathrm{E}[X_{n+k-1} - L(X_{n+k-1}|X_n, X_{n-1}, \ldots, X_{n-m-1})]^2 \\
&= \lim_{m\to\infty} \sigma_{k-1,m+1}^2 = \sigma_{k-1}^2. \qquad (2.6)
\end{aligned}$$

纯非决定性

- 决定性与非决定性取决于一步线性预报误差是否为零。
- 对非决定性序列，用$\{X_s, s \le n\}$预报X_{n+k}的误差会随k增大而增大。
- 记

$$\sigma_{k,m}^2 = \mathrm{E}[X_{n+k} - L(X_{n+k}|X_n, X_{n-1}, \ldots, X_{n-m+1})]^2$$

则$\sigma_{k,m}$也是m的单调递减函数，与n无关。可定义

$$\sigma_k^2 = \lim_{m\to\infty} \sigma_{k,m}^2$$

- 在极限意义下可以证明$\sigma_k^2 \ge \sigma_{k-1}^2$:

$$\begin{aligned}
\sigma_k^2 &= \lim_{m\to\infty} \mathrm{E}(X_{n+k} - L(X_{n+k}|X_n, X_{n-1}, \ldots, X_{n-m}))^2 \\
&= \lim_{m\to\infty} \mathrm{E}[X_{n+k-1} - L(X_{n+k-1}|X_{n-1}, X_{n-2}, \ldots, X_{n-1-m})]^2 \\
&\ge \lim_{m\to\infty} \mathrm{E}[X_{n+k-1} - L(X_{n+k-1}|X_n, X_{n-1}, \ldots, X_{n-m-1})]^2 \\
&= \lim_{m\to\infty} \sigma_{k-1,m+1}^2 = \sigma_{k-1}^2. \qquad (2.6)
\end{aligned}$$

纯非决定性II

- 注意上面证明中没有说明$\sigma^2_{k,m}$是k的增函数。
- 反例：AR(2)序列

$$X_t = \frac{1}{2}X_{t-2} + \varepsilon_t, \quad \varepsilon_t \sim \mathrm{WN}(0, \sigma^2)$$

- 平稳解为

$$X_t = \sum_{j=0}^{\infty} \left(\frac{1}{2}\right)^j \varepsilon_{t-2j}$$

-

$$\gamma_0 = \frac{4}{3}\sigma^2 \quad \gamma_1 = 0$$
$$\gamma_2 = \frac{2}{3}\sigma^2$$

-

$$L(X_t|X_{t-1}) = 0 \quad \sigma^2_{1,1} = \frac{4}{3}\sigma^2$$
$$L(X_t|X_{t-2}) = \frac{1}{2}X_{t-2} \quad \sigma^2_{2,1} = \sigma^2 < \sigma^2_{1,1}$$

纯非决定性II

- 注意上面证明中没有说明$\sigma^2_{k,m}$是k的增函数。
- 反例：AR(2)序列

$$X_t = \frac{1}{2}X_{t-2} + \varepsilon_t, \quad \varepsilon_t \sim \mathrm{WN}(0, \sigma^2)$$

- 平稳解为

$$X_t = \sum_{j=0}^{\infty} \left(\frac{1}{2}\right)^j \varepsilon_{t-2j}$$

-

$$\gamma_0 = \frac{4}{3}\sigma^2 \quad \gamma_1 = 0$$
$$\gamma_2 = \frac{2}{3}\sigma^2$$

-

$$L(X_t|X_{t-1}) = 0 \quad \sigma^2_{1,1} = \frac{4}{3}\sigma^2$$
$$L(X_t|X_{t-2}) = \frac{1}{2}X_{t-2} \quad \sigma^2_{2,1} = \sigma^2 < \sigma^2_{1,1}$$

纯非决定性II

- 注意上面证明中没有说明$\sigma^2_{k,m}$是k的增函数。
- 反例：AR(2)序列

$$X_t = \frac{1}{2}X_{t-2} + \varepsilon_t, \quad \varepsilon_t \sim \mathrm{WN}(0, \sigma^2)$$

- 平稳解为

$$X_t = \sum_{j=0}^{\infty} \left(\frac{1}{2}\right)^j \varepsilon_{t-2j}$$

-

$$\gamma_0 = \frac{4}{3}\sigma^2 \quad \gamma_1 = 0$$
$$\gamma_2 = \frac{2}{3}\sigma^2$$

-

$$L(X_t|X_{t-1}) = 0 \quad \sigma^2_{1,1} = \frac{4}{3}\sigma^2$$
$$L(X_t|X_{t-2}) = \frac{1}{2}X_{t-2} \quad \sigma^2_{2,1} = \sigma^2 < \sigma^2_{1,1}$$

纯非决定性II

- 注意上面证明中没有说明$\sigma^2_{k,m}$是k的增函数。
- 反例：AR(2)序列

$$X_t = \frac{1}{2}X_{t-2} + \varepsilon_t, \quad \varepsilon_t \sim \mathrm{WN}(0, \sigma^2)$$

- 平稳解为

$$X_t = \sum_{j=0}^{\infty} \left(\frac{1}{2}\right)^j \varepsilon_{t-2j}$$

- $$\gamma_0 = \frac{4}{3}\sigma^2 \quad \gamma_1 = 0$$
$$\gamma_2 = \frac{2}{3}\sigma^2$$

- $$L(X_t|X_{t-1}) = 0 \quad \sigma^2_{1,1} = \frac{4}{3}\sigma^2$$
$$L(X_t|X_{t-2}) = \frac{1}{2}X_{t-2} \quad \sigma^2_{2,1} = \sigma^2 < \sigma^2_{1,1}$$

纯非决定性II

- 注意上面证明中没有说明$\sigma^2_{k,m}$是k的增函数。
- 反例：AR(2)序列

$$X_t = \frac{1}{2}X_{t-2} + \varepsilon_t, \quad \varepsilon_t \sim \text{WN}(0, \sigma^2)$$

- 平稳解为

$$X_t = \sum_{j=0}^{\infty} \left(\frac{1}{2}\right)^j \varepsilon_{t-2j}$$

-

$$\begin{aligned} \gamma_0 =& \frac{4}{3}\sigma^2 \quad \gamma_1 = 0 \\ \gamma_2 =& \frac{2}{3}\sigma^2 \end{aligned}$$

-

$$\begin{aligned} L(X_t | X_{t-1}) =& 0 \quad \sigma^2_{1,1} = \frac{4}{3}\sigma^2 \\ L(X_t | X_{t-2}) =& \frac{1}{2}X_{t-2} \quad \sigma^2_{2,1} = \sigma^2 < \sigma^2_{1,1} \end{aligned}$$

纯非决定性III

- 由最佳线性预测定义知

$$\begin{aligned}\sigma^2_{k,m} =& \mathrm{E}[X_{n+k} - L(X_{n+k}|X_n, X_{n-1}, \ldots, X_{n-m+1})]^2 \\ \le& \mathrm{E}[X_{n+k} - 0]^2 = \gamma_0\end{aligned}$$

  所以$\sigma^2_k \le \gamma_0$。
- $k \to \infty$时如果$\sigma^2_k \to \gamma_0$ 则最佳线性预测与用平均值0预测效果相同，没有作用。
- **定义2.2** 设$\{X_t\}$是非决定性的平稳序列. 如果$\lim_{k\to\infty} \sigma^2_k = \gamma_0$, 则称$\{X_t\}$是纯非决定性的.
- 纯非决定性的平稳列不能作长期预报。
- 非决定性但不是纯非决定性的平稳列作长期预报是有意义的; 当然，决定性序列可以精确地长期预报。

纯非决定性III

- 由最佳线性预测定义知

$$\begin{aligned}\sigma_{k,m}^2 =& \mathrm{E}[X_{n+k} - L(X_{n+k}|X_n, X_{n-1}, \ldots, X_{n-m+1})]^2 \\ \le& \mathrm{E}[X_{n+k} - 0]^2 = \gamma_0\end{aligned}$$

  所以$\sigma_k^2 \le \gamma_0$。
- $k \to \infty$时如果$\sigma_k^2 \to \gamma_0$ 则最佳线性预测与用平均值0预测效果相同，没有作用。
- 定义2.2　设$\{X_t\}$是非决定性的平稳序列．如果$\lim_{k\to\infty}\sigma_k^2 = \gamma_0$，则称$\{X_t\}$是纯非决定性的．
- 纯非决定性的平稳列不能作长期预报。
- 非决定性但不是纯非决定性的平稳列作长期预报是有意义的；当然，决定性序列可以精确地长期预报。

纯非决定性III

- 由最佳线性预测定义知

$$\begin{aligned}\sigma_{k,m}^2 =&\mathrm{E}[X_{n+k} - L(X_{n+k}|X_n, X_{n-1}, \ldots, X_{n-m+1})]^2\\ \le&\mathrm{E}[X_{n+k} - 0]^2 = \gamma_0\end{aligned}$$

  所以$\sigma_k^2 \le \gamma_0$。
- $k \to \infty$时如果$\sigma_k^2 \to \gamma_0$ 则最佳线性预测与用平均值0预测效果相同，没有作用。
- **定义2.2** 设$\{X_t\}$是非决定性的平稳序列. 如果$\lim_{k\to\infty} \sigma_k^2 = \gamma_0$, 则称$\{X_t\}$是纯非决定性的.
- 纯非决定性的平稳列不能作长期预报。
- 非决定性但不是纯非决定性的平稳列作长期预报是有意义的; 当然，决定性序列可以精确地长期预报。

纯非决定性III

- 由最佳线性预测定义知

$$\begin{aligned}\sigma_{k,m}^2 =&\mathrm{E}[X_{n+k} - L(X_{n+k}|X_n, X_{n-1}, \ldots, X_{n-m+1})]^2 \\ \leq&\mathrm{E}[X_{n+k} - 0]^2 = \gamma_0\end{aligned}$$

  所以$\sigma_k^2 \leq \gamma_0$。
- $k \to \infty$时如果$\sigma_k^2 \to \gamma_0$ 则最佳线性预测与用平均值0预测效果相同，没有作用。
- **定义2.2** 设$\{X_t\}$是非决定性的平稳序列. 如果$\lim_{k\to\infty} \sigma_k^2 = \gamma_0$, 则称$\{X_t\}$是纯非决定性的.
- 纯非决定性的平稳列不能作长期预报。
- 非决定性但不是纯非决定性的平稳列作长期预报是有意义的; 当然，决定性序列可以精确地长期预报。

纯非决定性III

- 由最佳线性预测定义知

$$
\begin{aligned}
\sigma_{k,m}^2 =& \mathrm{E}[X_{n+k} - L(X_{n+k}|X_n, X_{n-1}, \ldots, X_{n-m+1})]^2 \\
\le& \mathrm{E}[X_{n+k} - 0]^2 = \gamma_0
\end{aligned}
$$

  所以$\sigma_k^2 \le \gamma_0$。
- $k \to \infty$时如果$\sigma_k^2 \to \gamma_0$ 则最佳线性预测与用平均值0预测效果相同，没有作用。
- **定义2.2** 设$\{X_t\}$是非决定性的平稳序列. 如果$\lim_{k\to\infty} \sigma_k^2 = \gamma_0$, 则称$\{X_t\}$是纯非决定性的.
- 纯非决定性的平稳列不能作长期预报。
- 非决定性但不是纯非决定性的平稳列作长期预报是有意义的; 当然，决定性序列可以精确地长期预报。

纯非决定性IV

- 对纯非决定性的平稳序列，有如下的结果：

$$\lim_{k\to\infty}\lim_{m\to\infty} \mathrm{E}[L(X_{n+k}|X_n, X_{n-1}, \ldots, X_{n-m+1})]^2 = 0. \tag{2.7}$$

- 实际上，记$\hat{X}_{n+k,m} = L(X_{n+k}|X_n, X_{n-1}, \ldots, X_{n-m+1})$. 由投影的正交性得

$$\sigma^2_{k,m} = E(X_{n+k} - \hat{X}_{n+k,m})^2 = EX^2_{n+k} - E\hat{X}^2_{n+k,m}.$$

于是得到

$$\lim_{k\to\infty}\lim_{m\to\infty} \mathrm{E}\hat{X}^2_{n+k,m} = \lim_{k\to\infty}\lim_{m\to\infty} (\gamma_0 - \sigma^2_{k,m}) = \gamma_0 - \gamma_0 = 0. \tag{2.8}$$

- 从(2.7)也可看出，对于纯非决定性的平稳序列做长期或超长期预测是不合适的.

纯非决定性IV

- 对纯非决定性的平稳序列，有如下的结果:

$$\lim_{k\to\infty}\lim_{m\to\infty}\mathrm{E}[L(X_{n+k}|X_n,X_{n-1},\ldots,X_{n-m+1})]^2=0. \tag{2.7}$$

- 实际上，记$\hat{X}_{n+k,m}=L(X_{n+k}|X_n,X_{n-1},\ldots,X_{n-m+1})$. 由投影的正交性得

$$\sigma_{k,m}^2=E(X_{n+k}-\hat{X}_{n+k,m})^2=EX_{n+k}^2-E\hat{X}_{n+k,m}^2.$$

于是得到

$$\lim_{k\to\infty}\lim_{m\to\infty}\mathrm{E}\hat{X}_{n+k,m}^2=\lim_{k\to\infty}\lim_{m\to\infty}(\gamma_0-\sigma_{k,m}^2)=\gamma_0-\gamma_0=0. \tag{2.8}$$

- 从(2.7)也可看出，对于纯非决定性的平稳序列做长期或超长期预测是不合适的.

纯非决定性IV

- 对纯非决定性的平稳序列，有如下的结果:

$$\lim_{k\to\infty}\lim_{m\to\infty}\mathrm{E}[L(X_{n+k}|X_n,X_{n-1},\ldots,X_{n-m+1})]^2=0. \tag{2.7}$$

- 实际上，记$\hat{X}_{n+k,m}=L(X_{n+k}|X_n,X_{n-1},\ldots,X_{n-m+1})$. 由投影的正交性得

$$\sigma_{k,m}^2=E(X_{n+k}-\hat{X}_{n+k,m})^2=EX_{n+k}^2-E\hat{X}_{n+k,m}^2.$$

于是得到

$$\lim_{k\to\infty}\lim_{m\to\infty}\mathrm{E}\hat{X}_{n+k,m}^2=\lim_{k\to\infty}\lim_{m\to\infty}(\gamma_0-\sigma_{k,m}^2)=\gamma_0-\gamma_0=0. \tag{2.8}$$

- 从(2.7)也可看出，对于纯非决定性的平稳序列做长期或超长期预测是不合适的.

线性闭包的等价定义

- **引理2.2** 设A为Hilbert空间H的子集，记L_A为A的所有有限线性组合构成的子集，记H_A为包含A的最小的闭子空间(称为由A生成的闭子空间)。则$\forall \xi \in H_A$，必存在$\xi_n \in L_A, n = 1, 2, \ldots$ 使得

$$\|\xi_n - \xi\| \to 0, \quad n \to \infty.$$

- 证明： 易见L_A为H的子线性空间，H_A为H的子Hilbert空间，$L_A \subset H_A$。记$\bar{L}_A$为L_A及L_A中元素的极限构成的集合，因$L_A \subset H_A$及H_A是闭集可知$\bar{L}_A \subset H_A$。
另一方面，可以证明$\bar{L}_A$是闭子空间，由H_A的定义及$A \subset L_A$得$H_A \subset \bar{L}_A$。

线性闭包的等价定义

- **引理2.2** 设A为Hilbert空间H的子集，记L_A为A的所有有限线性组合构成的子集，记H_A为包含A的最小的闭子空间(称为由A生成的闭子空间)。则$\forall \xi \in H_A$，必存在$\xi_n \in L_A, n = 1, 2, \ldots$ 使得
$$\|\xi_n - \xi\| \to 0, \quad n \to \infty.$$
- **证明**: 易见L_A为H的子线性空间，H_A为H的子Hilbert空间，$L_A \subset H_A$。记$\bar{L}_A$为L_A及L_A中元素的极限构成的集合，因$L_A \subset H_A$及H_A是闭集可知$\bar{L}_A \subset H_A$。
另一方面，可以证明$\bar{L}_A$是闭子空间，由H_A的定义及$A \subset L_A$得$H_A \subset \bar{L}_A$。

线性预测的极限

- 记H_n为$X_n, X_{n-1}, \ldots$生成的闭子空间。
- $L(X_{n+k}|X_n, X_{n-1}, \ldots, X_{n-m+1})$当$m \to \infty$时为$L(X_{n+k}|H_n)$(定理2.4)。
- **定理2.3** 设$Y \in L^2$, $\xi \in H_n$, 则$\xi = L(Y|H_n)$的充分必要条件是

$$Y - \xi \perp X_j, \quad j = n, n-1, n-2, \ldots \qquad (*)$$

- 证明:
  - 必要性：由定理1.1得到$Y - \xi \perp H_n$所以有$(*)$。
  - 充分性：记$A = \{X_n, X_{n-1}, \ldots\}$，则由$(*)$可知$Y - \xi \perp L_A$。由引理2.2, 对$\eta \in H_n$有$\eta_m \in L_A$使$\eta_m \to \eta$，由内积的连续性可得

$$\mathrm{E}((Y-\xi)\eta) = \lim_{m\to\infty} \mathrm{E}((Y-\xi)\eta_m) = 0$$

即$Y - \xi \perp H_n$，由定理1.1即得$\xi = L(Y|H_n)$。

线性预测的极限

- 记H_n为$X_n, X_{n-1}, \ldots$生成的闭子空间。
- $L(X_{n+k}|X_n, X_{n-1}, \ldots, X_{n-m+1})$当$m \to \infty$时为$L(X_{n+k}|H_n)$(定理2.4)。
- **定理2.3** 设$Y \in L^2$, $\xi \in H_n$, 则$\xi = L(Y|H_n)$的充分必要条件是

$$Y - \xi \perp X_j, \quad j = n, n-1, n-2, \ldots \qquad (*)$$

- 证明:
  - 必要性：由定理1.1得到$Y - \xi \perp H_n$所以有(*)。
  - 充分性：记$A = \{X_n, X_{n-1}, \ldots\}$，则由(*)可知$Y - \xi \perp L_A$。由引理2.2, 对$\eta \in H_n$有$\eta_m \in L_A$使$\eta_m \to \eta$，由内积的连续性可得

$$\mathrm{E}((Y - \xi)\eta) = \lim_{m \to \infty} \mathrm{E}((Y - \xi)\eta_m) = 0$$

即$Y - \xi \perp H_n$，由定理1.1即得$\xi = L(Y|H_n)$。

线性预测的极限

- 记H_n为$X_n, X_{n-1}, \ldots$生成的闭子空间。
- $L(X_{n+k}|X_n, X_{n-1}, \ldots, X_{n-m+1})$当$m \to \infty$时为$L(X_{n+k}|H_n)$(定理2.4)。
- **定理2.3** 设$Y \in L^2$, $\xi \in H_n$, 则$\xi = L(Y|H_n)$的充分必要条件是

$$Y - \xi \perp X_j, \quad j = n, n-1, n-2, \ldots \qquad (*)$$

- 证明:
  - 必要性：由定理1.1得到$Y - \xi \perp H_n$所以有(*)。
  - 充分性：记$A = \{X_n, X_{n-1}, \ldots\}$，则由(*)可知$Y - \xi \perp L_A$。由引理2.2，对$\eta \in H_n$有$\eta_m \in L_A$使$\eta_m \to \eta$，由内积的连续性可得

$$\mathrm{E}((Y-\xi)\eta) = \lim_{m\to\infty} \mathrm{E}((Y-\xi)\eta_m) = 0$$

即$Y - \xi \perp H_n$，由定理1.1即得$\xi = L(Y|H_n)$。

线性预测的极限

- 记H_n为$X_n, X_{n-1}, \ldots$生成的闭子空间。
- $L(X_{n+k}|X_n, X_{n-1}, \ldots, X_{n-m+1})$当$m \to \infty$时为$L(X_{n+k}|H_n)$(定理2.4)。
- **定理2.3** 设$Y \in L^2$, $\xi \in H_n$, 则$\xi = L(Y|H_n)$的充分必要条件是

$$Y - \xi \perp X_j, \quad j = n, n-1, n-2, \ldots \qquad (*)$$

- **证明**:
  - 必要性：由定理1.1得到$Y - \xi \perp H_n$所以有(*)。
  - 充分性：记$A = \{X_n, X_{n-1}, \ldots\}$，则由(*)可知$Y - \xi \perp L_A$。由引理2.2, 对$\eta \in H_n$有$\eta_m \in L_A$使$\eta_m \to \eta$，由内积的连续性可得

$$\mathrm{E}((Y-\xi)\eta) = \lim_{m\to\infty} \mathrm{E}((Y-\xi)\eta_m) = 0$$

即$Y - \xi \perp H_n$，由定理1.1即得$\xi = L(Y|H_n)$。

线性预测的极限

- 记H_n为$X_n, X_{n-1}, \ldots$生成的闭子空间。
- $L(X_{n+k}|X_n, X_{n-1}, \ldots, X_{n-m+1})$当$m \to \infty$时为$L(X_{n+k}|H_n)$(定理2.4)。
- **定理2.3** 设$Y \in L^2$, $\xi \in H_n$, 则$\xi = L(Y|H_n)$的充分必要条件是

$$Y - \xi \perp X_j, \quad j = n, n-1, n-2, \ldots \qquad (*)$$

- **证明**:
  - 必要性：由定理1.1得到$Y - \xi \perp H_n$所以有(*)。
  - 充分性：记$A = \{X_n, X_{n-1}, \ldots\}$，则由(*)可知$Y - \xi \perp L_A$。由引理2.2, 对$\eta \in H_n$有$\eta_m \in L_A$使$\eta_m \to \eta$，由内积的连续性可得

$$\mathrm{E}((Y-\xi)\eta) = \lim_{m\to\infty} \mathrm{E}((Y-\xi)\eta_m) = 0$$

即$Y - \xi \perp H_n$，由定理1.1即得$\xi = L(Y|H_n)$。

线性预测的极限

- 记H_n为$X_n, X_{n-1}, \ldots$生成的闭子空间。
- $L(X_{n+k}|X_n, X_{n-1}, \ldots, X_{n-m+1})$当$m \to \infty$时为$L(X_{n+k}|H_n)$(定理2.4)。
- **定理2.3** 设$Y \in L^2$, $\xi \in H_n$, 则$\xi = L(Y|H_n)$的充分必要条件是

$$Y - \xi \perp X_j, \quad j = n, n-1, n-2, \ldots \qquad (*)$$

- **证明**:
  - 必要性：由定理1.1得到$Y - \xi \perp H_n$所以有(*)。
  - 充分性：记$A = \{X_n, X_{n-1}, \ldots\}$，则由(*)可知$Y - \xi \perp L_A$。由引理2.2,
    对$\eta \in H_n$有$\eta_m \in L_A$使$\eta_m \to \eta$，由内积的连续性可得

$$\mathrm{E}((Y - \xi)\eta) = \lim_{m\to\infty} \mathrm{E}((Y - \xi)\eta_m) = 0$$

即$Y - \xi \perp H_n$，由定理1.1即得$\xi = L(Y|H_n)$。

无穷维历史线性预报的意义

- H_n为$\{X_s, s \leq n\}$所张成的子Hilbert空间，$L(X_{n+k}|H_n)$是一个投影。下面的定理说明这个投影是有穷维最佳线性预测的极限。
- 定理2.4 设$\mathbf{X}_{n,m} = (X_n, X_{n-1}, \ldots, X_{n-m+1})^T$，当$m \to \infty$时

$$L(Y|\mathbf{X}_{n,m}) \xrightarrow{\text{m.s.}} \hat{Y} \triangleq L(Y|H_n) \qquad (2.11)$$

无穷维历史线性预报的意义

- H_n为$\{X_s, s \le n\}$所张成的子Hilbert空间，$L(X_{n+k}|H_n)$是一个投影。下面的定理说明这个投影是有穷维最佳线性预测的极限。
- **定理2.4** 设$\mathbf{X}_{n,m} = (X_n, X_{n-1}, \ldots, X_{n-m+1})^T$，当$m \to \infty$时

$$L(Y|\mathbf{X}_{n,m}) \xrightarrow{\text{m.s.}} \hat{Y} \triangleq L(Y|H_n) \qquad (2.11)$$

定理2.4证明

- **证明**： 记$\hat{Y}_m = L(Y|\mathbf{X}_{n,m})$。先证明$\{\hat{Y}_m\}$是$H_n$中基本列。显然$\hat{Y}_m \in H_n$，设当$m \to \infty$时

$$\eta_m^2 \triangleq \mathrm{E}(Y - \hat{Y}_m)^2 \to \eta^2 \qquad (注意单调性)$$

- 对$m, k \to \infty$，注意$\hat{Y}_m, \hat{Y}_{m+k}$都和$Y - \hat{Y}_{m+k}$正交，得

$$\begin{aligned}
&\|\hat{Y}_m - \hat{Y}_{m+k}\|^2 = \|\hat{Y}_m - Y + Y - \hat{Y}_{m+k}\|^2 \\
=&\|\hat{Y}_m - Y\|^2 + \|Y - \hat{Y}_{m+k}\|^2 + 2\langle \hat{Y}_m - Y, Y - \hat{Y}_{m+k}\rangle \\
=&\eta_m^2 + \eta_{m+k}^2 - 2\langle Y, Y - \hat{Y}_{m+k}\rangle \\
=&\eta_m^2 + \eta_{m+k}^2 - 2\langle Y - \hat{Y}_{m+k}, Y - \hat{Y}_{m+k}\rangle \\
=&\eta_m^2 + \eta_{m+k}^2 - 2\eta_{m+k}^2 \to 0
\end{aligned}$$

- 因此$\{\hat{Y}_m\}$是H_n的基本列，在H_n中存在唯一极限ξ。

定理2.4证明

- **证明**： 记$\hat{Y}_m = L(Y|\mathbf{X}_{n,m})$。先证明$\{\hat{Y}_m\}$是$H_n$中基本列。显然$\hat{Y}_m \in H_n$，设当$m \to \infty$时

$$\eta_m^2 \triangleq \mathrm{E}(Y - \hat{Y}_m)^2 \to \eta^2 \qquad \text{(注意单调性)}$$

- 对$m, k \to \infty$，注意$\hat{Y}_m, \hat{Y}_{m+k}$都和$Y - \hat{Y}_{m+k}$正交，得

$$\begin{aligned}
&\|\hat{Y}_m - \hat{Y}_{m+k}\|^2 = \|\hat{Y}_m - Y + Y - \hat{Y}_{m+k}\|^2 \\
=&\|\hat{Y}_m - Y\|^2 + \|Y - \hat{Y}_{m+k}\|^2 + 2\langle \hat{Y}_m - Y, Y - \hat{Y}_{m+k}\rangle \\
=&\eta_m^2 + \eta_{m+k}^2 - 2\langle Y, Y - \hat{Y}_{m+k}\rangle \\
=&\eta_m^2 + \eta_{m+k}^2 - 2\langle Y - \hat{Y}_{m+k}, Y - \hat{Y}_{m+k}\rangle \\
=&\eta_m^2 + \eta_{m+k}^2 - 2\eta_{m+k}^2 \to 0
\end{aligned}$$

- 因此$\{\hat{Y}_m\}$是H_n的基本列，在H_n中存在唯一极限ξ。

定理2.4证明

- **证明**：记$\hat{Y}_m = L(Y|\mathbf{X}_{n,m})$。先证明$\{\hat{Y}_m\}$是$H_n$中基本列。显然$\hat{Y}_m \in H_n$，设当$m \to \infty$时

$$\eta_m^2 \triangleq \mathrm{E}(Y - \hat{Y}_m)^2 \to \eta^2 \qquad \text{(注意单调性)}$$

- 对$m, k \to \infty$，注意$\hat{Y}_m, \hat{Y}_{m+k}$都和$Y - \hat{Y}_{m+k}$正交，得

$$\begin{aligned}
&\|\hat{Y}_m - \hat{Y}_{m+k}\|^2 = \|\hat{Y}_m - Y + Y - \hat{Y}_{m+k}\|^2 \\
=&\|\hat{Y}_m - Y\|^2 + \|Y - \hat{Y}_{m+k}\|^2 + 2\langle \hat{Y}_m - Y, Y - \hat{Y}_{m+k}\rangle \\
=&\eta_m^2 + \eta_{m+k}^2 - 2\langle Y, Y - \hat{Y}_{m+k}\rangle \\
=&\eta_m^2 + \eta_{m+k}^2 - 2\langle Y - \hat{Y}_{m+k}, Y - \hat{Y}_{m+k}\rangle \\
=&\eta_m^2 + \eta_{m+k}^2 - 2\eta_{m+k}^2 \to 0
\end{aligned}$$

- 因此$\{\hat{Y}_m\}$是H_n的基本列，在H_n中存在唯一极限ξ。

定理2.4证明II

- 由内积的连续性，对任何$X_s, s \le n$有
$$\langle X_s, Y-\xi\rangle = \lim_{m\to\infty}\langle X_s, Y-L(Y|\mathbf{X}_{n,m})\rangle = 0$$
- 由定理2.3得到$\xi = L(Y|H_n)$。 $\square$

定理2.4证明II

- 由内积的连续性，对任何$X_s, s \leq n$有
$$\langle X_s, Y - \xi \rangle = \lim_{m\to\infty} \langle X_s, Y - L(Y|\mathbf{X}_{n,m}) \rangle = 0$$
- 由定理2.3得到$\xi = L(Y|H_n)$。 $\square$

最佳线性预报的方差

- 由内积连续性，

$$\sigma_1^2 \triangleq \lim_{m\to\infty} \|X_{n+1} - L(X_{n+1}|\mathbf{X}_{n,m})\|^2$$
$$= \|X_{n+1} - L(X_{n+1}|H_n)\|^2 = \|X_1 - L(X_1|H_0)\|^2$$

- $\sigma_1^2 = 0 \Leftrightarrow X_1 = L(X_1|H_0)$，所以$\sigma_1^2 = 0 \Rightarrow X_1 \in H_0$。反之，如果$X_1 \in H_0$，则$E(X_1 - X_1)^2 = 0$最小所以$X_1 = L(X_1|H_0)$。即$\sigma_1^2 = 0 \Leftrightarrow X_1 \in H_0$。这时$\{X_t\}$是决定性序列。
- 类似地，

$$\sigma_k^2 \triangleq \lim_{m\to\infty} \|X_{n+k} - L(X_{n+k}|\mathbf{X}_{n,m})\|^2$$
$$= \|X_{n+k} - L(X_{n+k}|H_n)\|^2 = \|X_k - L(X_k|H_0)\|^2$$

最佳线性预报的方差

- 由内积连续性，

$$\sigma_1^2 \triangleq \lim_{m\to\infty} \|X_{n+1} - L(X_{n+1}|\mathbf{X}_{n,m})\|^2$$
$$= \|X_{n+1} - L(X_{n+1}|H_n)\|^2 = \|X_1 - L(X_1|H_0)\|^2$$

- $\sigma_1^2 = 0 \Leftrightarrow X_1 = L(X_1|H_0)$，所以$\sigma_1^2 = 0 \Rightarrow X_1 \in H_0$。反之，如果$X_1 \in H_0$，则$E(X_1 - X_1)^2 = 0$最小所以$X_1 = L(X_1|H_0)$。即$\sigma_1^2 = 0 \Leftrightarrow X_1 \in H_0$。这时$\{X_t\}$是决定性序列。
- 类似地，

$$\sigma_k^2 \triangleq \lim_{m\to\infty} \|X_{n+k} - L(X_{n+k}|\mathbf{X}_{n,m})\|^2$$
$$= \|X_{n+k} - L(X_{n+k}|H_n)\|^2 = \|X_k - L(X_k|H_0)\|^2$$

最佳线性预报的方差

- 由内积连续性，

$$\begin{aligned}\sigma_1^2 &\triangleq \lim_{m\to\infty} \|X_{n+1} - L(X_{n+1}|\mathbf{X}_{n,m})\|^2 \\ &= \|X_{n+1} - L(X_{n+1}|H_n)\|^2 = \|X_1 - L(X_1|H_0)\|^2\end{aligned}$$

- $\sigma_1^2 = 0 \Leftrightarrow X_1 = L(X_1|H_0)$，所以$\sigma_1^2 = 0 \Rightarrow X_1 \in H_0$。反之，如果$X_1 \in H_0$，则$E(X_1 - X_1)^2 = 0$最小所以$X_1 = L(X_1|H_0)$。即$\sigma_1^2 = 0 \Leftrightarrow X_1 \in H_0$。这时$\{X_t\}$是决定性序列。
- 类似地，

$$\begin{aligned}\sigma_k^2 &\triangleq \lim_{m\to\infty} \|X_{n+k} - L(X_{n+k}|\mathbf{X}_{n,m})\|^2 \\ &= \|X_{n+k} - L(X_{n+k}|H_n)\|^2 = \|X_k - L(X_k|H_0)\|^2\end{aligned}$$

最佳线性预报的方差II

- **定理2.5** 设$\{X_t\}$是零均值平稳列，
  1. $\{X_t\}$是决定性序列当且仅当对某个n有
  $$X_{n+1} \in H_n; \tag{2.14}$$
  并且如果(2.14)对某个n成立则对所有n成立，这时$H_n = H_{n-1}, \forall n \in \mathbb{Z}$。
  2. $\{X_t\}$是纯非决定性的当且仅当对某个n，有
  $$\sigma_k^2 = \|X_{n+k} - L(X_{n+k}|H_n)\|^2 \to \gamma_0, \quad k \to \infty \tag{2.15}$$
  并且如果(2.15)对某个n成立则对所有n成立。

最佳线性预报的方差II

- **定理2.5** 设$\{X_t\}$是零均值平稳列，
  1. $\{X_t\}$是决定性序列当且仅当对某个n有
  $$X_{n+1} \in H_n; \tag{2.14}$$
  并且如果(2.14)对某个n成立则对所有n成立，这时$H_n = H_{n-1}, \forall n \in \mathbb{Z}$。
  2. $\{X_t\}$是纯非决定性的当且仅当对某个n，有
  $$\sigma_k^2 = \|X_{n+k} - L(X_{n+k}|H_n)\|^2 \to \gamma_0, \quad k \to \infty \tag{2.15}$$
  并且如果(2.15)对某个n成立则对所有n成立。

最佳线性预报的方差II

- **定理2.5** 设$\{X_t\}$是零均值平稳列，
  1. $\{X_t\}$是决定性序列当且仅当对某个n有
  $$X_{n+1} \in H_n; \tag{2.14}$$
  并且如果(2.14)对某个n成立则对所有n成立，这时$H_n = H_{n-1}, \forall n \in \mathbb{Z}$。
  2. $\{X_t\}$是纯非决定性的当且仅当对某个n，有
  $$\sigma_k^2 = \|X_{n+k} - L(X_{n+k}|H_n)\|^2 \to \gamma_0, \quad k \to \infty \tag{2.15}$$
  并且如果(2.15)对某个n成立则对所有n成立。

Wold表示定理

- **定理2.6**(Wold表示定理): 任一非决定性的零均值平稳列可以表示成

$$X_t = \sum_{j=0}^{\infty} a_j \varepsilon_{t-j} + V_t, \quad t \in \mathbb{Z} \qquad (2.16)$$

其中

- (1) $\varepsilon_t = X_t - L(X_t | X_{t-1}, X_{t-2}, \ldots)$ 是零均值白噪声，满足

$$\mathrm{E}\varepsilon_t^2 = \sigma^2 > 0, \quad a_0 = 1$$

$$a_j = \mathrm{E}(X_t \varepsilon_{t-j})/\sigma^2,$$

$$\sum_{j=0}^{\infty} a_j^2 < \infty$$

Wold表示定理

- **定理2.6**(Wold表示定理): 任一非决定性的零均值平稳列可以表示成

$$X_t = \sum_{j=0}^{\infty} a_j \varepsilon_{t-j} + V_t, \quad t \in \mathbb{Z} \qquad (2.16)$$

其中

- (1) $\varepsilon_t = X_t - L(X_t|X_{t-1}, X_{t-2}, \ldots)$ 是零均值白噪声，满足

$$\mathrm{E}\varepsilon_t^2 = \sigma^2 > 0, \quad a_0 = 1$$

$$a_j = \mathrm{E}(X_t \varepsilon_{t-j})/\sigma^2,$$

$$\sum_{j=0}^{\infty} a_j^2 < \infty$$

Wold表示定理II

- (续)
  - (2) $\{U_t = \sum_{j=0}^{\infty} a_j \varepsilon_{t-j},\ t \in \mathbb{Z}\}$ 是与$\{V_t\}$正交的平稳列;
  - (3) 定义$H_\varepsilon(t) = \overline{\mathrm{sp}}\{\varepsilon_s : s \le t\}$, $H_U(t) = \overline{\mathrm{sp}}\{U_s : s \le t\}$, 则$\forall t$
    $$H_U(t) = H_\varepsilon(t)$$
  - (4) $\{U_t\}$是纯非决定性的平稳序列，有谱密度
    $$f(\lambda) = \frac{\sigma^2}{2\pi}\left|\sum_{j=0}^{\infty} a_j e^{ij\lambda}\right|^2$$
  - (5) $\{V_t\}$是决定性的平稳序列。对任何$t, k \in \mathbb{Z}$, $V_t \in H_{t-k}$。

Wold表示定理II

- (续)
  - (2) $\{U_t = \sum_{j=0}^{\infty} a_j \varepsilon_{t-j},\ t \in \mathbb{Z}\}$ 是与$\{V_t\}$正交的平稳序列；
  - (3) 定义$H_\varepsilon(t) = \overline{\text{sp}}\{\varepsilon_s : s \le t\}$, $H_U(t) = \overline{\text{sp}}\{U_s : s \le t\}$, 则$\forall t$

$$H_U(t) = H_\varepsilon(t)$$

  - (4) $\{U_t\}$是纯非决定性的平稳序列，有谱密度

$$f(\lambda) = \frac{\sigma^2}{2\pi} \left| \sum_{j=0}^{\infty} a_j e^{ij\lambda} \right|^2$$

  - (5) $\{V_t\}$是决定性的平稳序列。对任何$t, k \in \mathbb{Z}$，$V_t \in H_{t-k}$。

Wold表示定理II

- (续)
  - (2) $\{U_t = \sum_{j=0}^{\infty} a_j \varepsilon_{t-j},\ t \in \mathbb{Z}\}$ 是与$\{V_t\}$正交的平稳列；
  - (3) 定义$H_\varepsilon(t) = \overline{\text{sp}}\{\varepsilon_s : s \le t\}$, $H_U(t) = \overline{\text{sp}}\{U_s : s \le t\}$, 则$\forall t$

$$H_U(t) = H_\varepsilon(t)$$

  - (4) $\{U_t\}$是纯非决定性的平稳序列，有谱密度

$$f(\lambda) = \frac{\sigma^2}{2\pi} \left| \sum_{j=0}^{\infty} a_j e^{ij\lambda} \right|^2$$

  - (5) $\{V_t\}$是决定性的平稳序列。对任何$t, k \in \mathbb{Z}$，$V_t \in H_{t-k}$。

Wold表示定理II

- (续)
  - (2) $\{U_t = \sum_{j=0}^{\infty} a_j \varepsilon_{t-j}, \ t \in \mathbb{Z}\}$ 是与$\{V_t\}$正交的平稳列；
  - (3) 定义$H_\varepsilon(t) = \overline{\text{sp}}\{\varepsilon_s : s \leq t\}$, $H_U(t) = \overline{\text{sp}}\{U_s : s \leq t\}$, 则$\forall t$
$$H_U(t) = H_\varepsilon(t)$$
  - (4) $\{U_t\}$是纯非决定性的平稳序列，有谱密度
$$f(\lambda) = \frac{\sigma^2}{2\pi} \left| \sum_{j=0}^{\infty} a_j e^{ij\lambda} \right|^2$$
  - (5) $\{V_t\}$是决定性的平稳序列。对任何$t, k \in \mathbb{Z}$，$V_t \in H_{t-k}$。

Wold表示定理II

- (续)
  - (2) $\{U_t = \sum_{j=0}^{\infty} a_j \varepsilon_{t-j},\ t \in \mathbb{Z}\}$ 是与$\{V_t\}$正交的平稳列；
  - (3) 定义$H_\varepsilon(t) = \overline{\text{sp}}\{\varepsilon_s : s \le t\}$, $H_U(t) = \overline{\text{sp}}\{U_s : s \le t\}$, 则$\forall t$

$$H_U(t) = H_\varepsilon(t)$$

  - (4) $\{U_t\}$是纯非决定性的平稳序列，有谱密度

$$f(\lambda) = \frac{\sigma^2}{2\pi}\left|\sum_{j=0}^{\infty} a_j e^{ij\lambda}\right|^2$$

  - (5) $\{V_t\}$是决定性的平稳序列。对任何$t, k \in \mathbb{Z}$，$V_t \in H_{t-k}$。

Wold表示定理的意义

- **定义2.3** 在Wold表示定理中
  - (1) 称(2.16)是$\{X_t\}$的Wold表示；
  - (2) 称$\{U_t\}$是$\{X_t\}$的纯非决定性部分，称$\{V_t\}$是$\{X_t\}$的决定性部分；
  - (3) 称$\{a_j\}$是$\{X_t\}$的Wold系数；
  - (4) 称一步预测误差$\varepsilon_t = X_t - L(X_t|X_{t-1}, X_{t-2}, \ldots)$为$\{X_t\}$的新息序列;
  - (5) 称$\sigma^2 = \mathrm{E}\varepsilon_t^2$为一步(线性)预测的均方误差。
- 这里新息的意思是完全不能被历史线性预测。
- 由Wold定理可知任何纯非决定性平稳序列可以表达为新息的单边滑动和。
- 事实上，任何白噪声的单边滑动和(系数平方可和)一定是纯非决定性的，但其中的白噪声不一定恰好是新息。

Wold表示定理的意义

- **定义2.3** 在Wold表示定理中
  - (1) 称(2.16)是$\{X_t\}$的Wold表示；
  - (2) 称$\{U_t\}$是$\{X_t\}$的纯非决定性部分，称$\{V_t\}$是$\{X_t\}$的决定性部分；
  - (3) 称$\{a_j\}$是$\{X_t\}$的Wold系数；
  - (4) 称一步预测误差$\varepsilon_t = X_t - L(X_t|X_{t-1}, X_{t-2}, \dots)$为$\{X_t\}$的新息序列；
  - (5) 称$\sigma^2 = \mathrm{E}\varepsilon_t^2$为一步(线性)预测的均方误差。
- 这里新息的意思是完全不能被历史线性预测。
- 由Wold定理可知任何纯非决定性平稳序列可以表达为新息的单边滑动和。
- 事实上，任何白噪声的单边滑动和(系数平方可和)一定是纯非决定性的，但其中的白噪声不一定恰好是新息。

Wold表示定理的意义

- **定义2.3** 在Wold表示定理中
  - (1) 称(2.16)是$\{X_t\}$的Wold表示；
  - (2) 称$\{U_t\}$是$\{X_t\}$的纯非决定性部分，称$\{V_t\}$是$\{X_t\}$的决定性部分；
  - (3) 称$\{a_j\}$是$\{X_t\}$的Wold系数；
  - (4) 称一步预测误差$\varepsilon_t = X_t - L(X_t|X_{t-1}, X_{t-2}, \ldots)$为$\{X_t\}$的新息序列；
  - (5) 称$\sigma^2 = \mathrm{E}\varepsilon_t^2$为一步(线性)预测的均方误差。
- 这里新息的意思是完全不能被历史线性预测。
- 由Wold定理可知任何纯非决定性平稳序列可以表达为新息的单边滑动和。
- 事实上，任何白噪声的单边滑动和(系数平方可和)一定是纯非决定性的，但其中的白噪声不一定恰好是新息。

Wold表示定理的意义

- **定义2.3** 在Wold表示定理中
  - (1) 称(2.16)是$\{X_t\}$的**Wold表示**；
  - (2) 称$\{U_t\}$是$\{X_t\}$的**纯非决定性部分**，称$\{V_t\}$是$\{X_t\}$的**决定性部分**；
  - (3) 称$\{a_j\}$是$\{X_t\}$的**Wold系数**；
  - (4) 称一步预测误差$\varepsilon_t = X_t - L(X_t|X_{t-1}, X_{t-2}, \ldots)$为$\{X_t\}$的**新息序列**;
  - (5) 称$\sigma^2 = \mathrm{E}\varepsilon_t^2$为**一步(线性)预测的均方误差**。
- 这里新息的意思是完全不能被历史线性预测。
- 由Wold定理可知任何纯非决定性平稳序列可以表达为新息的单边滑动和。
- 事实上，任何白噪声的单边滑动和(系数平方可和)一定是纯非决定性的，但其中的白噪声不一定恰好是新息。

Wold表示定理的意义

- **定义2.3** 在Wold表示定理中
  - (1) 称(2.16)是$\{X_t\}$的Wold表示；
  - (2) 称$\{U_t\}$是$\{X_t\}$的纯非决定性部分，称$\{V_t\}$是$\{X_t\}$的决定性部分；
  - (3) 称$\{a_j\}$是$\{X_t\}$的Wold系数；
  - (4) 称一步预测误差$\varepsilon_t = X_t - L(X_t|X_{t-1}, X_{t-2}, \dots)$为$\{X_t\}$的新息序列；
  - (5) 称$\sigma^2 = \mathrm{E}\varepsilon_t^2$为一步(线性)预测的均方误差。
- 这里新息的意思是完全不能被历史线性预测。
- 由Wold定理可知任何纯非决定性平稳序列可以表达为新息的单边滑动和。
- 事实上，任何白噪声的单边滑动和(系数平方可和)一定是纯非决定性的，但其中的白噪声不一定恰好是新息。

Wold表示定理的意义

- **定义2.3** 在Wold表示定理中
  - (1) 称(2.16)是$\{X_t\}$的Wold表示；
  - (2) 称$\{U_t\}$是$\{X_t\}$的纯非决定性部分，称$\{V_t\}$是$\{X_t\}$的决定性部分；
  - (3) 称$\{a_j\}$是$\{X_t\}$的Wold系数；
  - (4) 称一步预测误差$\varepsilon_t = X_t - L(X_t|X_{t-1}, X_{t-2}, \dots)$为$\{X_t\}$的新息序列;
  - (5) 称$\sigma^2 = \mathrm{E}\varepsilon_t^2$为一步(线性)预测的均方误差。
- 这里新息的意思是完全不能被历史线性预测。
- 由Wold定理可知任何纯非决定性平稳序列可以表达为新息的单边滑动和。
- 事实上，任何白噪声的单边滑动和(系数平方可和)一定是纯非决定性的，但其中的白噪声不一定恰好是新息。

Wold表示定理的意义

- **定义2.3** 在Wold表示定理中
  - (1) 称(2.16)是$\{X_t\}$的Wold表示；
  - (2) 称$\{U_t\}$是$\{X_t\}$的纯非决定性部分，称$\{V_t\}$是$\{X_t\}$的决定性部分；
  - (3) 称$\{a_j\}$是$\{X_t\}$的Wold系数；
  - (4) 称一步预测误差$\varepsilon_t = X_t - L(X_t | X_{t-1}, X_{t-2}, \dots)$为$\{X_t\}$的新息序列;
  - (5) 称$\sigma^2 = \mathrm{E}\varepsilon_t^2$为一步(线性)预测的均方误差。
- 这里新息的意思是完全不能被历史线性预测。
- 由Wold定理可知任何纯非决定性平稳序列可以表达为新息的单边滑动和。
- 事实上，任何白噪声的单边滑动和(系数平方可和)一定是纯非决定性的，但其中的白噪声不一定恰好是新息。

Wold表示定理的意义

- **定义2.3**　在Wold表示定理中
  - (1) 称(2.16)是$\{X_t\}$的Wold表示；
  - (2) 称$\{U_t\}$是$\{X_t\}$的纯非决定性部分，称$\{V_t\}$是$\{X_t\}$的决定性部分；
  - (3) 称$\{a_j\}$是$\{X_t\}$的Wold系数；
  - (4) 称一步预测误差$\varepsilon_t = X_t - L(X_t|X_{t-1}, X_{t-2}, \dots)$为$\{X_t\}$的新息序列;
  - (5) 称$\sigma^2 = \mathrm{E}\varepsilon_t^2$为一步(线性)预测的均方误差。
- 这里新息的意思是完全不能被历史线性预测。
- 由Wold定理可知任何纯非决定性平稳序列可以表达为新息的单边滑动和。
- 事实上，任何白噪声的单边滑动和(系数平方可和)一定是纯非决定性的，但其中的白噪声不一定恰好是新息。

Wold表示定理的意义

- **定义2.3** 在Wold表示定理中
  - (1) 称(2.16)是$\{X_t\}$的**Wold表示**；
  - (2) 称$\{U_t\}$是$\{X_t\}$的**纯非决定性部分**，称$\{V_t\}$是$\{X_t\}$的**决定性部分**；
  - (3) 称$\{a_j\}$是$\{X_t\}$的**Wold系数**；
  - (4) 称一步预测误差$\varepsilon_t = X_t - L(X_t | X_{t-1}, X_{t-2}, \dots)$为$\{X_t\}$的**新息序列**;
  - (5) 称$\sigma^2 = \mathrm{E}\varepsilon_t^2$为**一步(线性)预测的均方误差**。
- 这里新息的意思是完全不能被历史线性预测。
- 由Wold定理可知任何纯非决定性平稳序列可以表达为新息的单边滑动和。
- 事实上，任何白噪声的单边滑动和(系数平方可和)一定是纯非决定性的，但其中的白噪声不一定恰好是新息。

ARMA序列的Wold表示

- **例2.3** 设$\{X_t\}$是ARMA(p,q)序列，模型方程为

$$A(\mathscr{B})X_t = B(\mathscr{B})\varepsilon_t, \quad t \in \mathbb{Z},$$

- 设$A^{-1}(z)B(z)$有Taylor展开式

$$\Psi(z) = A^{-1}(z)B(z) = \sum_{j=0}^{\infty} \psi_j z^j, \quad |z| \le 1$$

- 则

$$X_t = \sum_{j=0}^{\infty} \psi_j \varepsilon_{t-j}, \quad t \in \mathbb{Z}. \qquad (2.19)$$

- 下面证明$\{X_t\}$是纯非决定性的平稳序列，(2.19)是$\{X_t\}$的Wold表示，$\{\varepsilon_t\}$是$\{X_t\}$的新息序列，$\{\psi_j\}$是$\{X_t\}$的Wold系数。

ARMA序列的Wold表示

- **例2.3** 设$\{X_t\}$是ARMA(p,q)序列，模型方程为

$$A(\mathscr{B})X_t = B(\mathscr{B})\varepsilon_t, \quad t \in \mathbb{Z},$$

- 设$A^{-1}(z)B(z)$有Taylor展开式

$$\Psi(z) = A^{-1}(z)B(z) = \sum_{j=0}^{\infty} \psi_j z^j, \quad |z| \le 1$$

- 则

$$X_t = \sum_{j=0}^{\infty} \psi_j \varepsilon_{t-j}, \quad t \in \mathbb{Z}. \qquad (2.19)$$

- 下面证明$\{X_t\}$是纯非决定性的平稳序列，(2.19)是$\{X_t\}$的Wold表示，$\{\varepsilon_t\}$是$\{X_t\}$的新息序列，$\{\psi_j\}$是$\{X_t\}$的Wold系数。

ARMA序列的Wold表示

- **例2.3** 设$\{X_t\}$是ARMA(p,q)序列，模型方程为

$$A(\mathscr{B})X_t = B(\mathscr{B})\varepsilon_t, \quad t \in \mathbb{Z},$$

- 设$A^{-1}(z)B(z)$有Taylor展开式

$$\Psi(z) = A^{-1}(z)B(z) = \sum_{j=0}^{\infty} \psi_j z^j, \quad |z| \le 1$$

- 则

$$X_t = \sum_{j=0}^{\infty} \psi_j \varepsilon_{t-j}, \quad t \in \mathbb{Z}. \tag{2.19}$$

- 下面证明$\{X_t\}$是纯非决定性的平稳序列，(2.19)是$\{X_t\}$的Wold表示，$\{\varepsilon_t\}$是$\{X_t\}$的新息序列，$\{\psi_j\}$是$\{X_t\}$的Wold系数。

ARMA序列的Wold表示

- **例2.3** 设$\{X_t\}$是ARMA(p,q)序列，模型方程为

$$A(\mathscr{B})X_t = B(\mathscr{B})\varepsilon_t, \quad t\in\mathbb{Z},$$

- 设$A^{-1}(z)B(z)$有Taylor展开式

$$\Psi(z) = A^{-1}(z)B(z) = \sum_{j=0}^{\infty}\psi_j z^j, \quad |z|\le 1$$

- 则

$$X_t = \sum_{j=0}^{\infty}\psi_j\varepsilon_{t-j}, \quad t\in\mathbb{Z}. \tag{2.19}$$

- 下面证明$\{X_t\}$是纯非决定性的平稳序列，(2.19)是$\{X_t\}$的Wold表示，$\{\varepsilon_t\}$是$\{X_t\}$的新息序列，$\{\psi_j\}$是$\{X_t\}$的Wold系数。

ARMA序列的Wold表示II

- 我们只对比较容易的可逆ARMA的情况证明。由(2.19)看出$X_t \in H_\varepsilon(t)$, 利用可逆性，$\varepsilon_t = B^{-1}(\mathscr{B})A(\mathscr{B})X_t \in H_t$, 所以$H_t = H_\varepsilon(t)$.
- 于是

$$X_t - \varepsilon_t = \sum_{j=1}^{\infty} \psi_j \varepsilon_{t-j} \in H_\varepsilon(t-1) = H_{t-1}$$

来证$X_t - \varepsilon_t = L(X_t|H_{t-1})$。只要证明$X_t - (X_t - \varepsilon_t) \perp H_{t-1}$.

ARMA序列的Wold表示II

- 我们只对比较容易的可逆ARMA的情况证明。由(2.19)看出$X_t \in H_\varepsilon(t)$, 利用可逆性，$\varepsilon_t = B^{-1}(\mathscr{B})A(\mathscr{B})X_t \in H_t$, 所以$H_t = H_\varepsilon(t)$.
- 于是

$$X_t - \varepsilon_t = \sum_{j=1}^{\infty} \psi_j \varepsilon_{t-j} \in H_\varepsilon(t-1) = H_{t-1}$$

来证$X_t - \varepsilon_t = L(X_t|H_{t-1})$。只要证明$X_t - (X_t - \varepsilon_t) \perp H_{t-1}$.

ARMA序列的Wold表示III

- 事实上，由于ε_t与$\varepsilon_{t-j}, j \geq 1$正交可知$\varepsilon_t \perp H_\varepsilon(t-1) = H_{t-1}$。故

$$\begin{aligned} X_t - \varepsilon_t =& L(X_t | H_{t-1}) \\ \varepsilon_t =& X_t - L(X_t | H_{t-1}) \end{aligned}$$

即$\{\varepsilon_t\}$是$\{X_t\}$的新息列，$\sigma^2 = \mathrm{E}\varepsilon_t^2$ 是一步预测均方误差，在(2.19)两边同乘以ε_{t-j}后取期望，利用内积的连续性可得

$$\psi_j = \langle X_t, \varepsilon_{t-j} \rangle / \sigma^2$$

即$\{\psi_j\}$是$\{X_t\}$的Wold系数列。

Wold定理的证明—I

- 取$\varepsilon_t = X_t - L(X_t|H_{t-1})$, $H_\varepsilon(t) = \overline{sp}\{\varepsilon_s: s \le t\}$, 易见$\varepsilon_t \in H_t$，由$H_t$的单调性可知$\varepsilon_t \in H_s, \forall t < s$，因此$H_\varepsilon(t) \subset H_t$。
- 来证明$\{\varepsilon_t\}$是白噪声。
- 由定理2.4

$$\begin{aligned} L(X_t|H_{t-1}) &= \lim_{m\to\infty} L(X_t|X_{t-1},\dots,X_{t-m}) \quad (L^2) \\ &= \lim_{m\to\infty} \mathbf{a}_m^T \mathbf{X}_{t-1,m}^T \end{aligned}$$

其中$\mathbf{a}_m$是m阶Y-W系数(预测方程的解)，不依赖于t，

- 由内积的连续性

$$\begin{aligned} \mathrm{E}\varepsilon_t^2 &= \lim_{m\to\infty} \|X_t - L(X_t|X_{t-1},\dots,X_{t-m})\|^2 \\ &= \lim_{m\to\infty} (\gamma_0 - \mathbf{a}_m^T \Gamma_m \mathbf{a}_m) \quad (\text{与}t\text{无关}) \\ &= \lim_{m\to\infty} \sigma_{1,m}^2 = \sigma^2 > 0 \quad (\text{由非决定性定义}) \end{aligned}$$

- 对$s > t$，$\varepsilon_s \perp H_{s-1} \supset H_t$所以$\varepsilon_s \perp \varepsilon_t, s > t$。即

$$\{\varepsilon_t\} \sim \mathrm{WN}(0, \sigma^2), \quad \sigma^2 > 0$$

Wold定理的证明—I

- 取$\varepsilon_t = X_t - L(X_t|H_{t-1})$, $H_\varepsilon(t) = \bar{sp}\{\varepsilon_s: s \le t\}$, 易见$\varepsilon_t \in H_t$，由$H_t$的单调性可知$\varepsilon_t \in H_s, \forall t < s$，因此$H_\varepsilon(t) \subset H_t$。
- 来证明$\{\varepsilon_t\}$是白噪声。
- 由定理2.4

$$L(X_t|H_{t-1}) = \lim_{m\to\infty} L(X_t|X_{t-1},\dots,X_{t-m}) \quad (L^2)$$
$$= \lim_{m\to\infty} \mathbf{a}_m^T \mathbf{X}_{t-1,m}^T$$

其中$\mathbf{a}_m$是m阶Y-W系数(预测方程的解)，不依赖于t，

- 由内积的连续性

$$\mathrm{E}\varepsilon_t^2 = \lim_{m\to\infty} \|X_t - L(X_t|X_{t-1},\dots,X_{t-m})\|^2$$
$$= \lim_{m\to\infty} (\gamma_0 - \mathbf{a}_m^T \Gamma_m \mathbf{a}_m) \quad (\text{与}t\text{无关})$$
$$= \lim_{m\to\infty} \sigma_{1,m}^2 = \sigma^2 > 0 \quad (\text{由非决定性定义})$$

- 对$s > t$，$\varepsilon_s \perp H_{s-1} \supset H_t$所以$\varepsilon_s \perp \varepsilon_t, s > t$。即

$$\{\varepsilon_t\} \sim \mathrm{WN}(0, \sigma^2), \quad \sigma^2 > 0$$

Wold定理的证明—I

- 取$\varepsilon_t = X_t - L(X_t|H_{t-1})$, $H_\varepsilon(t) = \bar{sp}\{\varepsilon_s: s \le t\}$, 易见$\varepsilon_t \in H_t$，由$H_t$的单调性可知$\varepsilon_t \in H_s, \forall t < s$，因此$H_\varepsilon(t) \subset H_t$。
- 来证明$\{\varepsilon_t\}$是白噪声。
- 由定理2.4

$$\begin{aligned} L(X_t|H_{t-1}) &= \lim_{m\to\infty} L(X_t|X_{t-1},\dots,X_{t-m}) \quad (L^2) \\ &= \lim_{m\to\infty} \mathbf{a}_m^T \mathbf{X}_{t-1,m}^T \end{aligned}$$

其中$\mathbf{a}_m$是m阶Y-W系数(预测方程的解)，不依赖于t，
- 由内积的连续性

$$\begin{aligned} \mathrm{E}\varepsilon_t^2 &= \lim_{m\to\infty} \|X_t - L(X_t|X_{t-1},\dots,X_{t-m})\|^2 \\ &= \lim_{m\to\infty} (\gamma_0 - \mathbf{a}_m^T \Gamma_m \mathbf{a}_m) \quad (\text{与}t\text{无关}) \\ &= \lim_{m\to\infty} \sigma_{1,m}^2 = \sigma^2 > 0 \quad (\text{由非决定性定义}) \end{aligned}$$

- 对$s > t$，$\varepsilon_s \perp H_{s-1} \supset H_t$所以$\varepsilon_s \perp \varepsilon_t, s > t$。即

$$\{\varepsilon_t\} \sim \mathrm{WN}(0, \sigma^2), \quad \sigma^2 > 0$$

Wold定理的证明—I

- 取$\varepsilon_t = X_t - L(X_t|H_{t-1})$, $H_\varepsilon(t) = \bar{sp}\{\varepsilon_s: \ s \le t\}$, 易见$\varepsilon_t \in H_t$，由$H_t$的单调性可知$\varepsilon_t \in H_s, \forall t < s$，因此$H_\varepsilon(t) \subset H_t$。
- 来证明$\{\varepsilon_t\}$是白噪声。
- 由定理2.4

$$\begin{aligned} L(X_t|H_{t-1}) &= \lim_{m\to\infty} L(X_t|X_{t-1},\dots,X_{t-m}) \quad (L^2) \\ &= \lim_{m\to\infty} \mathbf{a}_m^T \mathbf{X}_{t-1,m}^T \end{aligned}$$

  其中$\mathbf{a}_m$是m阶Y-W系数(预测方程的解)，不依赖于t，
- 由内积的连续性

$$\begin{aligned} \mathrm{E}\varepsilon_t^2 &= \lim_{m\to\infty} \|X_t - L(X_t|X_{t-1},\dots,X_{t-m})\|^2 \\ &= \lim_{m\to\infty} (\gamma_0 - \mathbf{a}_m^T \Gamma_m \mathbf{a}_m) \quad (\text{与}t\text{无关}) \\ &= \lim_{m\to\infty} \sigma_{1,m}^2 = \sigma^2 > 0 \quad (\text{由非决定性定义}) \end{aligned}$$

- 对$s > t$，$\varepsilon_s \perp H_{s-1} \supset H_t$所以$\varepsilon_s \perp \varepsilon_t, s > t$。即

$$\{\varepsilon_t\} \sim \mathrm{WN}(0, \sigma^2), \quad \sigma^2 > 0$$

Wold定理的证明—I

- 取$\varepsilon_t = X_t - L(X_t|H_{t-1})$, $H_\varepsilon(t) = \bar{sp}\{\varepsilon_s: \ s \le t\}$, 易见$\varepsilon_t \in H_t$，由$H_t$的单调性可知$\varepsilon_t \in H_s, \forall t < s$，因此$H_\varepsilon(t) \subset H_t$。
- 来证明$\{\varepsilon_t\}$是白噪声。
- 由定理2.4

$$
\begin{aligned}
L(X_t|H_{t-1}) &= \lim_{m\to\infty} L(X_t|X_{t-1},\dots,X_{t-m}) \quad (L^2) \\
&= \lim_{m\to\infty} \mathbf{a}_m^T \mathbf{X}_{t-1,m}^T
\end{aligned}
$$

其中$\mathbf{a}_m$是m阶Y-W系数(预测方程的解)，不依赖于t，
- 由内积的连续性

$$
\begin{aligned}
\mathrm{E}\varepsilon_t^2 &= \lim_{m\to\infty} \|X_t - L(X_t|X_{t-1},\dots,X_{t-m})\|^2 \\
&= \lim_{m\to\infty} (\gamma_0 - \mathbf{a}_m^T \Gamma_m \mathbf{a}_m) \quad (与t无关) \\
&= \lim_{m\to\infty} \sigma_{1,m}^2 = \sigma^2 > 0 \quad (由非决定性定义)
\end{aligned}
$$

- 对$s > t$，$\varepsilon_s \perp H_{s-1} \supset H_t$所以$\varepsilon_s \perp \varepsilon_t, s > t$。即

$$
\{\varepsilon_t\} \sim \mathrm{WN}(0,\sigma^2), \quad \sigma^2 > 0
$$

Wold定理的证明—II

- 定义$V_t = X_t - L(X_t|H_\varepsilon(t))$, 则$V_t \in H_t$。来证明$\{\varepsilon_t\}$和$\{V_t\}$正交。
- 由投影性质，$V_t \perp H_\varepsilon(t)$，即$V_t \perp \varepsilon_s, s \le t$。
- 当$s > t$时，注意$\varepsilon_s \perp H_{s-1} \supset H_t$，而$V_t \in H_t$所以$\varepsilon_s \perp V_t, s > t$。
- 于是$\{\varepsilon_t\}$和$\{V_t\}$正交。

Wold定理的证明—II

- 定义$V_t = X_t - L(X_t|H_\varepsilon(t))$, 则$V_t \in H_t$。来证明$\{\varepsilon_t\}$和$\{V_t\}$正交。
- 由投影性质，$V_t \perp H_\varepsilon(t)$，即$V_t \perp \varepsilon_s, s \leq t$。
- 当$s > t$时，注意$\varepsilon_s \perp H_{s-1} \supset H_t$，而$V_t \in H_t$所以$\varepsilon_s \perp V_t, s > t$。
- 于是$\{\varepsilon_t\}$和$\{V_t\}$正交。

Wold定理的证明—II

- 定义$V_t = X_t - L(X_t|H_\varepsilon(t))$, 则$V_t \in H_t$。来证明$\{\varepsilon_t\}$和$\{V_t\}$正交。
- 由投影性质，$V_t \perp H_\varepsilon(t)$，即$V_t \perp \varepsilon_s, s \leq t$。
- 当$s > t$时，注意$\varepsilon_s \perp H_{s-1} \supset H_t$，而$V_t \in H_t$所以$\varepsilon_s \perp V_t, s > t$。
- 于是$\{\varepsilon_t\}$和$\{V_t\}$正交。

Wold定理的证明—II

- 定义$V_t = X_t - L(X_t|H_\varepsilon(t))$, 则$V_t \in H_t$。来证明$\{\varepsilon_t\}$和$\{V_t\}$正交。
- 由投影性质，$V_t \perp H_\varepsilon(t)$，即$V_t \perp \varepsilon_s, s \leq t$。
- 当$s > t$时，注意$\varepsilon_s \perp H_{s-1} \supset H_t$，而$V_t \in H_t$所以$\varepsilon_s \perp V_t, s > t$。
- 于是$\{\varepsilon_t\}$和$\{V_t\}$正交。

Wold定理的证明—III

- 来证明$\{\varepsilon_t\}$和$\{X_t\}$平稳相关。
- 当$s > t$时$\varepsilon_s \perp H_t$所以$\langle \varepsilon_s, X_t \rangle = 0, s > t$。
- 当$s \le t$时，由定理2.4，

$$L(X_t|H_{t-1}) = \lim_{m\to\infty} L(X_t|\mathbf{X}_{t-1,m}) \quad (L^2)$$
$$= \lim_{m\to\infty} \mathbf{a}_m^T \mathbf{X}_{t-1,m}$$

- 由内积的连续性，$s \le t$时

$$\langle X_t, \varepsilon_s \rangle = \lim_{m\to\infty} \langle X_t, X_s - \mathbf{a}_m^T \mathbf{X}_{t-1,m} \rangle$$
$$= \gamma_{t-s} - \lim_{m\to\infty} (a_{m1}\gamma_{t-s+1} + \cdots + a_{mm}\gamma_{t-s+m})$$

只依赖于$t-s$。
- 所以$\{\varepsilon_t\}$和$\{X_t\}$平稳相关。

Wold定理的证明—III

- 来证明$\{\varepsilon_t\}$和$\{X_t\}$平稳相关。
- 当$s > t$时$\varepsilon_s \perp H_t$所以$\langle \varepsilon_s, X_t \rangle = 0, s > t$。
- 当$s \le t$时，由定理2.4，

$$\begin{aligned} L(X_t|H_{t-1}) &= \lim_{m\to\infty} L(X_t|\mathbf{X}_{t-1,m}) \quad (L^2) \\ &= \lim_{m\to\infty} \mathbf{a}_m^T \mathbf{X}_{t-1,m} \end{aligned}$$

- 由内积的连续性, $s \le t$时

$$\begin{aligned} \langle X_t, \varepsilon_s \rangle = & \lim_{m\to\infty} \langle X_t, X_s - \mathbf{a}_m^T \mathbf{X}_{t-1,m} \rangle \\ =& \gamma_{t-s} - \lim_{m\to\infty} (a_{m1}\gamma_{t-s+1} + \cdots + a_{mm}\gamma_{t-s+m}) \end{aligned}$$

只依赖于$t-s$。
- 所以$\{\varepsilon_t\}$和$\{X_t\}$平稳相关。

Wold定理的证明—III

- 来证明$\{\varepsilon_t\}$和$\{X_t\}$平稳相关。
- 当$s > t$时$\varepsilon_s \perp H_t$所以$\langle \varepsilon_s, X_t \rangle = 0, s > t$。
- 当$s \leq t$时，由定理2.4，

$$
\begin{aligned}
L(X_t | H_{t-1}) &= \lim_{m\to\infty} L(X_t | \mathbf{X}_{t-1,m}) \quad (L^2) \\
&= \lim_{m\to\infty} \mathbf{a}_m^T \mathbf{X}_{t-1,m}
\end{aligned}
$$

- 由内积的连续性, $s \leq t$时

$$
\begin{aligned}
\langle X_t, \varepsilon_s \rangle &= \lim_{m\to\infty} \langle X_t, X_s - \mathbf{a}_m^T \mathbf{X}_{t-1,m} \rangle \\
&= \gamma_{t-s} - \lim_{m\to\infty} (a_{m1}\gamma_{t-s+1} + \cdots + a_{mm}\gamma_{t-s+m})
\end{aligned}
$$

只依赖于$t-s$。
- 所以$\{\varepsilon_t\}$和$\{X_t\}$平稳相关。

Wold定理的证明—III

- 来证明$\{\varepsilon_t\}$和$\{X_t\}$平稳相关。
- 当$s > t$时$\varepsilon_s \perp H_t$所以$\langle \varepsilon_s, X_t \rangle = 0, s > t$。
- 当$s \leq t$时，由定理2.4，

$$\begin{aligned} L(X_t | H_{t-1}) &= \lim_{m\to\infty} L(X_t | \mathbf{X}_{t-1,m}) \quad (L^2) \\ &= \lim_{m\to\infty} \mathbf{a}_m^T \mathbf{X}_{t-1,m} \end{aligned}$$

- 由内积的连续性, $s \leq t$时

$$\begin{aligned} \langle X_t, \varepsilon_s \rangle = &\lim_{m\to\infty} \langle X_t, X_s - \mathbf{a}_m^T \mathbf{X}_{t-1,m} \rangle \\ =&\gamma_{t-s} - \lim_{m\to\infty} (a_{m1}\gamma_{t-s+1} + \cdots + a_{mm}\gamma_{t-s+m}) \end{aligned}$$

只依赖于$t-s$。

- 所以$\{\varepsilon_t\}$和$\{X_t\}$平稳相关。

Wold定理的证明—III

- 来证明$\{\varepsilon_t\}$和$\{X_t\}$平稳相关。
- 当$s > t$时$\varepsilon_s \perp H_t$所以$\langle \varepsilon_s, X_t \rangle = 0, s > t$。
- 当$s \leq t$时，由定理2.4，

$$\begin{aligned} L(X_t|H_{t-1}) =& \lim_{m\to\infty} L(X_t|\mathbf{X}_{t-1,m}) \quad (L^2) \\ =& \lim_{m\to\infty} \mathbf{a}_m^T \mathbf{X}_{t-1,m} \end{aligned}$$

- 由内积的连续性, $s \leq t$时

$$\begin{aligned} \langle X_t, \varepsilon_s \rangle =& \lim_{m\to\infty} \langle X_t, X_s - \mathbf{a}_m^T \mathbf{X}_{t-1,m} \rangle \\ =& \gamma_{t-s} - \lim_{m\to\infty} (a_{m1}\gamma_{t-s+1} + \cdots + a_{mm}\gamma_{t-s+m}) \end{aligned}$$

只依赖于$t-s$。
- 所以$\{\varepsilon_t\}$和$\{X_t\}$平稳相关。

Wold定理的证明—IV

- 由III的证明，可以定义

$$a_j = \langle X_t, \varepsilon_{t-j}\rangle/\sigma^2 \quad (j \geq 0)$$

a_j与t无关。且

$$\begin{aligned} a_0 &= \langle X_t, \varepsilon_t\rangle/\sigma^2 \\ &= \langle \varepsilon_t + L(X_t|H_{t-1}), \varepsilon_t\rangle/\sigma^2 \\ &= \langle \varepsilon_t, \varepsilon_t\rangle/\sigma^2 = 1 \end{aligned}$$

- 令$U_t = L(X_t|H_\varepsilon(t))$，
则$V_t = X_t - L(X_t|H_\varepsilon(t)) = X_t - U_t$，$X_t = U_t + V_t$。
- 来证明

$$U_t = \sum_{j=0}^{\infty} a_j \varepsilon_{t-j} \qquad (*)$$

Wold定理的证明—IV

- 由III的证明，可以定义

$$a_j = \langle X_t, \varepsilon_{t-j} \rangle / \sigma^2 \quad (j \geq 0)$$

a_j与t无关。且

$$\begin{aligned} a_0 &= \langle X_t, \varepsilon_t \rangle / \sigma^2 \\ &= \langle \varepsilon_t + L(X_t | H_{t-1}), \varepsilon_t \rangle / \sigma^2 \\ &= \langle \varepsilon_t, \varepsilon_t \rangle / \sigma^2 = 1 \end{aligned}$$

- 令$U_t = L(X_t | H_\varepsilon(t))$，
则$V_t = X_t - L(X_t | H_\varepsilon(t)) = X_t - U_t$，$X_t = U_t + V_t$。
- 来证明

$$U_t = \sum_{j=0}^{\infty} a_j \varepsilon_{t-j} \qquad (*)$$

Wold定理的证明—IV

- 由III的证明，可以定义

$$a_j = \langle X_t, \varepsilon_{t-j} \rangle / \sigma^2 \quad (j \geq 0)$$

a_j与t无关。且

$$\begin{aligned} a_0 &= \langle X_t, \varepsilon_t \rangle / \sigma^2 \\ &= \langle \varepsilon_t + L(X_t | H_{t-1}), \varepsilon_t \rangle / \sigma^2 \\ &= \langle \varepsilon_t, \varepsilon_t \rangle / \sigma^2 = 1 \end{aligned}$$

- 令$U_t = L(X_t | H_\varepsilon(t))$，
  则$V_t = X_t - L(X_t | H_\varepsilon(t)) = X_t - U_t$，$X_t = U_t + V_t$。
- 来证明

$$U_t = \sum_{j=0}^{\infty} a_j \varepsilon_{t-j} \qquad (*)$$

Wold定理的证明—IV(续)

- 定义$U_{t,n} = L(X_t | \varepsilon_t, \varepsilon_{t-1}, \dots, \varepsilon_{t-n})$.
  设$U_{t,n} = \sum_{j=0}^{n} b_j \varepsilon_{t-j}$，则
- 对$j = 0, 1, \dots, n$

$$
\begin{aligned}
\sigma^2 a_j =& \langle X_t, \varepsilon_{t-j} \rangle = \langle U_{t,n} + (X_t - U_{t,n}), \varepsilon_{t-j} \rangle \\
=& \langle U_{t,n}, \varepsilon_{t-j} \rangle = \sigma^2 b_j
\end{aligned}
$$

即$b_j = a_j$,

$$
U_{t,n} = L(X_t | \varepsilon_t, \varepsilon_{t-1}, \dots, \varepsilon_{t-n}) = \sum_{j=0}^{n} a_j \varepsilon_{t-j}
$$

Wold定理的证明—IV(续)

- 定义$U_{t,n} = L(X_t|\varepsilon_t, \varepsilon_{t-1}, \dots, \varepsilon_{t-n})$.
  设$U_{t,n} = \sum_{j=0}^{n} b_j \varepsilon_{t-j}$，则
- 对$j = 0, 1, \dots, n$

$$
\begin{aligned}
\sigma^2 a_j =& \langle X_t, \varepsilon_{t-j} \rangle = \langle U_{t,n} + (X_t - U_{t,n}), \varepsilon_{t-j} \rangle \\
=& \langle U_{t,n}, \varepsilon_{t-j} \rangle = \sigma^2 b_j
\end{aligned}
$$

即$b_j = a_j$，

$$
U_{t,n} = L(X_t|\varepsilon_t, \varepsilon_{t-1}, \dots, \varepsilon_{t-n}) = \sum_{j=0}^{n} a_j \varepsilon_{t-j}
$$

Wold定理的证明—IV(续2)

- 注意$U_{t,n}$是X_t的投影所以$\|U_{t,n}\|^2 \leq \|X_t\|^2 = \gamma_0$，所以

$$\|U_{t,n}\|^2 = \sigma^2 \sum_{j=0}^{n} a_j^2 \leq \gamma_0 < \infty$$

- 故$\sum_{j=0}^{\infty} a_j^2 < \infty$，$\sum_{j=0}^{n} a_j \varepsilon_{t-j}$均方收敛到$\sum_{j=0}^{\infty} a_j \varepsilon_{t-j}$。
- 由定理2.4知$U_{t,n}$均方收敛到$U_t = L(X_t|H_\varepsilon(t))$，所以

$$U_t = \sum_{j=0}^{\infty} a_j \varepsilon_{t-j} \qquad (*)$$

- $\{U_t\}$是线性平稳列，其谱密度立即可得(结论(4)的第二部分)。
- 由于$\{\varepsilon_t\}$与$\{V_t\}$正交所以$V_s \perp H_\varepsilon(t), \forall s, t \in \mathbb{Z}$，而$U_t \in H_\varepsilon(t)$所以$V_s \perp U_t$，$\{V_t\}$与$\{U_t\}$正交。
- 由$\{U_t\}$和$\{V_t\}$正交，$X_t = U_t + V_t$，$\{X_t\}$和$\{U_t\}$平稳可知$V_t = X_t - U_t$也是平稳列。
- 至此定理的(1)(2)已证明。

Wold定理的证明—IV(续2)

- 注意$U_{t,n}$是X_t的投影所以$\|U_{t,n}\|^2 \leq \|X_t\|^2 = \gamma_0$，所以

$$\|U_{t,n}\|^2 = \sigma^2 \sum_{j=0}^{n} a_j^2 \leq \gamma_0 < \infty$$

- 故$\sum_{j=0}^{\infty} a_j^2 < \infty$，$\sum_{j=0}^{n} a_j \varepsilon_{t-j}$均方收敛到$\sum_{j=0}^{\infty} a_j \varepsilon_{t-j}$。
- 由定理2.4知$U_{t,n}$均方收敛到$U_t = L(X_t | H_\varepsilon(t))$，所以

$$U_t = \sum_{j=0}^{\infty} a_j \varepsilon_{t-j} \qquad (*)$$

- $\{U_t\}$是线性平稳列，其谱密度立即可得(结论(4)的第二部分)。
- 由于$\{\varepsilon_t\}$与$\{V_t\}$正交所以$V_s \perp H_\varepsilon(t), \forall s, t \in \mathbb{Z}$，而$U_t \in H_\varepsilon(t)$所以$V_s \perp U_t$，$\{V_t\}$与$\{U_t\}$正交。
- 由$\{U_t\}$和$\{V_t\}$正交，$X_t = U_t + V_t$，$\{X_t\}$和$\{U_t\}$平稳可知$V_t = X_t - U_t$也是平稳列。
- 至此定理的(1)(2)已证明。

Wold定理的证明—IV(续2)

- 注意$U_{t,n}$是X_t的投影所以$\|U_{t,n}\|^2 \leq \|X_t\|^2 = \gamma_0$，所以

$$\|U_{t,n}\|^2 = \sigma^2 \sum_{j=0}^{n} a_j^2 \leq \gamma_0 < \infty$$

- 故$\sum_{j=0}^{\infty} a_j^2 < \infty$，$\sum_{j=0}^{n} a_j \varepsilon_{t-j}$均方收敛到$\sum_{j=0}^{\infty} a_j \varepsilon_{t-j}$。
- 由定理2.4知$U_{t,n}$均方收敛到$U_t = L(X_t | H_\varepsilon(t))$，所以

$$U_t = \sum_{j=0}^{\infty} a_j \varepsilon_{t-j} \qquad (*)$$

- $\{U_t\}$是线性平稳列，其谱密度立即可得(结论(4)的第二部分)。
- 由于$\{\varepsilon_t\}$与$\{V_t\}$正交所以$V_s \perp H_\varepsilon(t), \forall s, t \in \mathbb{Z}$，而$U_t \in H_\varepsilon(t)$所以$V_s \perp U_t$，$\{V_t\}$与$\{U_t\}$正交。
- 由$\{U_t\}$和$\{V_t\}$正交，$X_t = U_t + V_t$，$\{X_t\}$和$\{U_t\}$平稳可知$V_t = X_t - U_t$也是平稳列。
- 至此定理的(1)(2)已证明。

Wold定理的证明—IV(续2)

- 注意$U_{t,n}$是X_t的投影所以$\|U_{t,n}\|^2 \le \|X_t\|^2 = \gamma_0$，所以

$$\|U_{t,n}\|^2 = \sigma^2 \sum_{j=0}^{n} a_j^2 \le \gamma_0 < \infty$$

- 故$\sum_{j=0}^{\infty} a_j^2 < \infty$，$\sum_{j=0}^{n} a_j \varepsilon_{t-j}$均方收敛到$\sum_{j=0}^{\infty} a_j \varepsilon_{t-j}$。
- 由定理2.4知$U_{t,n}$均方收敛到$U_t = L(X_t | H_\varepsilon(t))$，所以

$$U_t = \sum_{j=0}^{\infty} a_j \varepsilon_{t-j} \qquad (*)$$

- $\{U_t\}$是线性平稳列，其谱密度立即可得(结论(4)的第二部分)。
- 由于$\{\varepsilon_t\}$与$\{V_t\}$正交所以$V_s \perp H_\varepsilon(t), \forall s,t \in \mathbb{Z}$，而$U_t \in H_\varepsilon(t)$所以$V_s \perp U_t$, $\{V_t\}$与$\{U_t\}$正交。
- 由$\{U_t\}$和$\{V_t\}$正交，$X_t = U_t + V_t$, $\{X_t\}$和$\{U_t\}$平稳可知$V_t = X_t - U_t$也是平稳列。
- 至此定理的(1)(2)已证明。

Wold定理的证明—IV(续2)

- 注意$U_{t,n}$是X_t的投影所以$\|U_{t,n}\|^2 \leq \|X_t\|^2 = \gamma_0$，所以
$$\|U_{t,n}\|^2 = \sigma^2 \sum_{j=0}^{n} a_j^2 \leq \gamma_0 < \infty$$
- 故$\sum_{j=0}^{\infty} a_j^2 < \infty$，$\sum_{j=0}^{n} a_j \varepsilon_{t-j}$均方收敛到$\sum_{j=0}^{\infty} a_j \varepsilon_{t-j}$。
- 由定理2.4知$U_{t,n}$均方收敛到$U_t = L(X_t | H_\varepsilon(t))$，所以
$$U_t = \sum_{j=0}^{\infty} a_j \varepsilon_{t-j} \qquad (*)$$
- $\{U_t\}$是线性平稳列，其谱密度立即可得(结论(4)的第二部分)。
- 由于$\{\varepsilon_t\}$与$\{V_t\}$正交所以$V_s \perp H_\varepsilon(t), \forall s, t \in \mathbb{Z}$，而$U_t \in H_\varepsilon(t)$所以$V_s \perp U_t$, $\{V_t\}$与$\{U_t\}$正交。
- 由$\{U_t\}$和$\{V_t\}$正交，$X_t = U_t + V_t$, $\{X_t\}$和$\{U_t\}$平稳可知$V_t = X_t - U_t$也是平稳列。
- 至此定理的(1)(2)已证明。

Wold定理的证明—IV(续2)

- 注意$U_{t,n}$是X_t的投影所以$\|U_{t,n}\|^2 \leq \|X_t\|^2 = \gamma_0$，所以

$$\|U_{t,n}\|^2 = \sigma^2 \sum_{j=0}^{n} a_j^2 \leq \gamma_0 < \infty$$

- 故$\sum_{j=0}^{\infty} a_j^2 < \infty$，$\sum_{j=0}^{n} a_j \varepsilon_{t-j}$均方收敛到$\sum_{j=0}^{\infty} a_j \varepsilon_{t-j}$。
- 由定理2.4知$U_{t,n}$均方收敛到$U_t = L(X_t|H_\varepsilon(t))$，所以

$$U_t = \sum_{j=0}^{\infty} a_j \varepsilon_{t-j} \qquad (*)$$

- $\{U_t\}$是线性平稳列，其谱密度立即可得(结论(4)的第二部分)。
- 由于$\{\varepsilon_t\}$与$\{V_t\}$正交所以$V_s \perp H_\varepsilon(t), \forall s, t \in \mathbb{Z}$，而$U_t \in H_\varepsilon(t)$所以$V_s \perp U_t$, $\{V_t\}$与$\{U_t\}$正交。
- 由$\{U_t\}$和$\{V_t\}$正交，$X_t = U_t + V_t$, $\{X_t\}$和$\{U_t\}$平稳可知$V_t = X_t - U_t$也是平稳列。
- 至此定理的(1)(2)已证明。

Wold定理的证明—IV(续2)

- 注意$U_{t,n}$是X_t的投影所以$\|U_{t,n}\|^2 \leq \|X_t\|^2 = \gamma_0$，所以
$$\|U_{t,n}\|^2 = \sigma^2 \sum_{j=0}^{n} a_j^2 \leq \gamma_0 < \infty$$
- 故$\sum_{j=0}^{\infty} a_j^2 < \infty$，$\sum_{j=0}^{n} a_j \varepsilon_{t-j}$均方收敛到$\sum_{j=0}^{\infty} a_j \varepsilon_{t-j}$。
- 由定理2.4知$U_{t,n}$均方收敛到$U_t = L(X_t|H_\varepsilon(t))$，所以
$$U_t = \sum_{j=0}^{\infty} a_j \varepsilon_{t-j} \qquad (*)$$
- $\{U_t\}$是线性平稳列，其谱密度立即可得(结论(4)的第二部分)。
- 由于$\{\varepsilon_t\}$与$\{V_t\}$正交所以$V_s \perp H_\varepsilon(t), \forall s, t \in \mathbb{Z}$，而$U_t \in H_\varepsilon(t)$所以$V_s \perp U_t$, $\{V_t\}$与$\{U_t\}$正交。
- 由$\{U_t\}$和$\{V_t\}$正交，$X_t = U_t + V_t$, $\{X_t\}$和$\{U_t\}$平稳可知$V_t = X_t - U_t$也是平稳列。
- 至此定理的(1)(2)已证明。

Wold定理的证明—V

- 来证明第(3)条结论。定义$H_U(t) = \overline{\text{sp}}\{U_s: s \le t\}$, 来证明$H_U(t) = H_\varepsilon(t)$。
- 显然$U_t \in H_\varepsilon(t)$所以$H_U(t) \subset H_\varepsilon(t)$。只要证明$H_\varepsilon(t) \subset H_U(t)$。
- 注意$\varepsilon_t \in H_t \subset \overline{\text{sp}}\{U_s, V_s: s \le t\}$, 由引理2.2，存在$\xi_m \in L\{V_t, V_{t-1}, \dots, V_{t-m}\}$, $\eta_m \in L\{U_t, U_{t-1}, \dots, U_{t-m}\}$, 使

$$\|\xi_m + \eta_m - \varepsilon_t\|^2 \to 0$$

- 但前面已证明$\{V_t\}$与$\{\varepsilon_t\}$正交，也与$\{U_t\}$正交，所以

$$\begin{aligned}&\|\xi_m + \eta_m - \varepsilon_t\|^2\\ =&\|\xi_m\|^2 + \|\eta_m - \varepsilon_t\|^2\\ \ge&\|\eta_m - \varepsilon_t\|^2\end{aligned}$$

- 令$m \to \infty$得$\|\eta_m - \varepsilon_t\|^2 \to 0$，由引理2.2知$\varepsilon_t \in H_U(t)$。
- 所以$H_\varepsilon(t) \subset H_U(t)$, $H_\varepsilon(t) = H_U(t)$。结论(3)证毕。

Wold定理的证明—V

- 来证明第(3)条结论。定义$H_U(t) = \overline{\text{sp}}\{U_s : s \le t\}$, 来证明$H_U(t) = H_\varepsilon(t)$。
- 显然$U_t \in H_\varepsilon(t)$所以$H_U(t) \subset H_\varepsilon(t)$。只要证明$H_\varepsilon(t) \subset H_U(t)$。
- 注意$\varepsilon_t \in H_t \subset \overline{\text{sp}}\{U_s, V_s : s \le t\}$, 由引理2.2，存在$\xi_m \in L\{V_t, V_{t-1}, \dots, V_{t-m}\}$, $\eta_m \in L\{U_t, U_{t-1}, \dots, U_{t-m}\}$, 使

$$\|\xi_m + \eta_m - \varepsilon_t\|^2 \to 0$$

- 但前面已证明$\{V_t\}$与$\{\varepsilon_t\}$正交，也与$\{U_t\}$正交，所以

$$\begin{aligned}&\|\xi_m + \eta_m - \varepsilon_t\|^2\\ =&\|\xi_m\|^2 + \|\eta_m - \varepsilon_t\|^2\\ \ge&\|\eta_m - \varepsilon_t\|^2\end{aligned}$$

- 令$m \to \infty$得$\|\eta_m - \varepsilon_t\|^2 \to 0$，由引理2.2知$\varepsilon_t \in H_U(t)$。
- 所以$H_\varepsilon(t) \subset H_U(t)$, $H_\varepsilon(t) = H_U(t)$。结论(3)证毕。

Wold定理的证明—V

- 来证明第(3)条结论。定义$H_U(t) = \overline{\text{sp}}\{U_s: s \le t\}$，来证明$H_U(t) = H_\varepsilon(t)$。
- 显然$U_t \in H_\varepsilon(t)$所以$H_U(t) \subset H_\varepsilon(t)$。只要证明$H_\varepsilon(t) \subset H_U(t)$。
- 注意$\varepsilon_t \in H_t \subset \overline{\text{sp}}\{U_s, V_s: s \le t\}$，由引理2.2，存在$\xi_m \in L\{V_t, V_{t-1}, \dots, V_{t-m}\}$，$\eta_m \in L\{U_t, U_{t-1}, \dots, U_{t-m}\}$，使

$$\|\xi_m + \eta_m - \varepsilon_t\|^2 \to 0$$

- 但前面已证明$\{V_t\}$与$\{\varepsilon_t\}$正交，也与$\{U_t\}$正交，所以

$$\begin{aligned}&\|\xi_m + \eta_m - \varepsilon_t\|^2\\ =&\|\xi_m\|^2 + \|\eta_m - \varepsilon_t\|^2\\ \ge&\|\eta_m - \varepsilon_t\|^2\end{aligned}$$

- 令$m \to \infty$得$\|\eta_m - \varepsilon_t\|^2 \to 0$，由引理2.2知$\varepsilon_t \in H_U(t)$。
- 所以$H_\varepsilon(t) \subset H_U(t)$，$H_\varepsilon(t) = H_U(t)$。结论(3)证毕。

Wold定理的证明—V

- 来证明第(3)条结论。定义$H_U(t) = \overline{\text{sp}}\{U_s: \ s \le t\}$，来证明$H_U(t) = H_\varepsilon(t)$。
- 显然$U_t \in H_\varepsilon(t)$所以$H_U(t) \subset H_\varepsilon(t)$。只要证明$H_\varepsilon(t) \subset H_U(t)$。
- 注意$\varepsilon_t \in H_t \subset \overline{\text{sp}}\{U_s, V_s: \ s \le t\}$，由引理2.2，存在$\xi_m \in L\{V_t, V_{t-1}, \dots, V_{t-m}\}$，$\eta_m \in L\{U_t, U_{t-1}, \dots, U_{t-m}\}$，使

$$\|\xi_m + \eta_m - \varepsilon_t\|^2 \to 0$$

- 但前面已证明$\{V_t\}$与$\{\varepsilon_t\}$正交，也与$\{U_t\}$正交，所以

$$\begin{aligned} &\|\xi_m + \eta_m - \varepsilon_t\|^2 \\ =&\|\xi_m\|^2 + \|\eta_m - \varepsilon_t\|^2 \\ \ge&\|\eta_m - \varepsilon_t\|^2 \end{aligned}$$

- 令$m \to \infty$得$\|\eta_m - \varepsilon_t\|^2 \to 0$，由引理2.2知$\varepsilon_t \in H_U(t)$。
- 所以$H_\varepsilon(t) \subset H_U(t)$，$H_\varepsilon(t) = H_U(t)$。结论(3)证毕。

Wold定理的证明—V

- 来证明第(3)条结论。定义$H_U(t)=\overline{\text{sp}}\{U_s:\ s\le t\}$，来证明$H_U(t)=H_\varepsilon(t)$。
- 显然$U_t\in H_\varepsilon(t)$所以$H_U(t)\subset H_\varepsilon(t)$。只要证明$H_\varepsilon(t)\subset H_U(t)$。
- 注意$\varepsilon_t\in H_t\subset\overline{\text{sp}}\{U_s,V_s:\ s\le t\}$，由引理2.2，存在$\xi_m\in L\{V_t,V_{t-1},\dots,V_{t-m}\}$，$\eta_m\in L\{U_t,U_{t-1},\dots,U_{t-m}\}$，使

$$\|\xi_m+\eta_m-\varepsilon_t\|^2\to 0$$

- 但前面已证明$\{V_t\}$与$\{\varepsilon_t\}$正交，也与$\{U_t\}$正交，所以

$$\begin{aligned}&\|\xi_m+\eta_m-\varepsilon_t\|^2\\=&\|\xi_m\|^2+\|\eta_m-\varepsilon_t\|^2\\\ge&\|\eta_m-\varepsilon_t\|^2\end{aligned}$$

- 令$m\to\infty$得$\|\eta_m-\varepsilon_t\|^2\to 0$，由引理2.2知$\varepsilon_t\in H_U(t)$。
- 所以$H_\varepsilon(t)\subset H_U(t)$，$H_\varepsilon(t)=H_U(t)$。结论(3)证毕。

Wold定理的证明—V

- 来证明第(3)条结论。定义$H_U(t) = \overline{\text{sp}}\{U_s: \ s \le t\}$, 来证明$H_U(t) = H_\varepsilon(t)$。
- 显然$U_t \in H_\varepsilon(t)$所以$H_U(t) \subset H_\varepsilon(t)$。只要证明$H_\varepsilon(t) \subset H_U(t)$。
- 注意$\varepsilon_t \in H_t \subset \overline{\text{sp}}\{U_s, V_s: \ s \le t\}$, 由引理2.2，存在$\xi_m \in L\{V_t, V_{t-1}, \dots, V_{t-m}\}$, $\eta_m \in L\{U_t, U_{t-1}, \dots, U_{t-m}\}$, 使

$$\|\xi_m + \eta_m - \varepsilon_t\|^2 \to 0$$

- 但前面已证明$\{V_t\}$与$\{\varepsilon_t\}$正交，也与$\{U_t\}$正交，所以

$$\begin{aligned}
&\|\xi_m + \eta_m - \varepsilon_t\|^2 \\
=&\|\xi_m\|^2 + \|\eta_m - \varepsilon_t\|^2 \\
\ge&\|\eta_m - \varepsilon_t\|^2
\end{aligned}$$

- 令$m \to \infty$得$\|\eta_m - \varepsilon_t\|^2 \to 0$，由引理2.2知$\varepsilon_t \in H_U(t)$。
- 所以$H_\varepsilon(t) \subset H_U(t)$, $H_\varepsilon(t) = H_U(t)$。结论(3)证毕。

Wold定理的证明—VI

- 来证明$\{U_t\}$是纯非决定性的(结论(4))。
- 利用定理1.2(4)(5)

$$
\begin{aligned}
L(U_{t+k}|H_U(t)) &= L(U_{t+k}|H_\varepsilon(t)) \\
&= L\left[\sum_{j=0}^{k-1} a_j\varepsilon_{t+k-j} + \sum_{j=k}^{\infty} a_j\varepsilon_{t+k-j}|H_\varepsilon(t)\right] \\
&= \sum_{j=k}^{\infty} a_j\varepsilon_{t+k-j}
\end{aligned}
$$

- 于是

$$
\begin{aligned}
\|U_{t+k} - L(U_{t+k}|H_U(t))\|^2 &= \|\sum_{j=0}^{k-1} a_j\varepsilon_{t+k-j}\|^2 \\
&= \sigma^2\sum_{j=0}^{k-1} a_j^2 \to \sigma^2\sum_{j=0}^{\infty} a_j^2 = \mathrm{E}U_t^2
\end{aligned}
$$

- 按定义可知$\{U_t\}$为纯非决定性的。
- 注意：这个证明对一般单边线性序列不适用。

Wold定理的证明—VI

- 来证明$\{U_t\}$是纯非决定性的(结论(4))。
- 利用定理1.2(4)(5)

$$
\begin{aligned}
L(U_{t+k}|H_U(t)) &= L(U_{t+k}|H_\varepsilon(t)) \\
&= L\left[\sum_{j=0}^{k-1} a_j\varepsilon_{t+k-j} + \sum_{j=k}^{\infty} a_j\varepsilon_{t+k-j}|H_\varepsilon(t)\right] \\
&= \sum_{j=k}^{\infty} a_j\varepsilon_{t+k-j}
\end{aligned}
$$

- 于是

$$
\begin{aligned}
\|U_{t+k} - L(U_{t+k}|H_U(t))\|^2 &= \|\sum_{j=0}^{k-1} a_j\varepsilon_{t+k-j}\|^2 \\
&= \sigma^2\sum_{j=0}^{k-1} a_j^2 \to \sigma^2\sum_{j=0}^{\infty} a_j^2 = \mathrm{E}U_t^2
\end{aligned}
$$

- 按定义可知$\{U_t\}$为纯非决定性的。
- 注意：这个证明对一般单边线性序列不适用。

Wold定理的证明—VI

- 来证明$\{U_t\}$是纯非决定性的(结论(4))。
- 利用定理1.2(4)(5)

$$
\begin{aligned}
L(U_{t+k}|H_U(t)) &= L(U_{t+k}|H_\varepsilon(t)) \\
=&L\left[\sum_{j=0}^{k-1} a_j\varepsilon_{t+k-j} + \sum_{j=k}^{\infty} a_j\varepsilon_{t+k-j}|H_\varepsilon(t)\right] \\
=&\sum_{j=k}^{\infty} a_j\varepsilon_{t+k-j}
\end{aligned}
$$

- 于是

$$
\begin{aligned}
\|U_{t+k} - L(U_{t+k}|H_U(t))\|^2 &= \|\sum_{j=0}^{k-1} a_j\varepsilon_{t+k-j}\|^2 \\
=&\sigma^2\sum_{j=0}^{k-1} a_j^2 \to \sigma^2\sum_{j=0}^{\infty} a_j^2 = \mathrm{E}U_t^2
\end{aligned}
$$

- 按定义可知$\{U_t\}$为纯非决定性的。
- 注意：这个证明对一般单边线性序列不适用。

Wold定理的证明—VI

- 来证明$\{U_t\}$是纯非决定性的(结论(4))。
- 利用定理1.2(4)(5)

$$
\begin{aligned}
&L(U_{t+k}|H_U(t)) = L(U_{t+k}|H_\varepsilon(t)) \\
=&L\left[\sum_{j=0}^{k-1} a_j\varepsilon_{t+k-j} + \sum_{j=k}^{\infty} a_j\varepsilon_{t+k-j}|H_\varepsilon(t)\right] \\
=&\sum_{j=k}^{\infty} a_j\varepsilon_{t+k-j}
\end{aligned}
$$

- 于是

$$
\begin{aligned}
&\|U_{t+k} - L(U_{t+k}|H_U(t))\|^2 = \|\sum_{j=0}^{k-1} a_j\varepsilon_{t+k-j}\|^2 \\
=&\sigma^2\sum_{j=0}^{k-1} a_j^2 \to \sigma^2\sum_{j=0}^{\infty} a_j^2 = \mathrm{E}U_t^2
\end{aligned}
$$

- 按定义可知$\{U_t\}$为纯非决定性的。
- 注意：这个证明对一般单边线性序列不适用。

Wold定理的证明—VI

- 来证明$\{U_t\}$是纯非决定性的(结论(4))。
- 利用定理1.2(4)(5)

$$
\begin{aligned}
L(U_{t+k}|H_U(t)) &= L(U_{t+k}|H_\varepsilon(t)) \\
=&L\left[\sum_{j=0}^{k-1} a_j\varepsilon_{t+k-j} + \sum_{j=k}^{\infty} a_j\varepsilon_{t+k-j}|H_\varepsilon(t)\right] \\
=&\sum_{j=k}^{\infty} a_j\varepsilon_{t+k-j}
\end{aligned}
$$

- 于是

$$
\|U_{t+k} - L(U_{t+k}|H_U(t))\|^2 = \|\sum_{j=0}^{k-1} a_j\varepsilon_{t+k-j}\|^2
$$

$$
=\sigma^2\sum_{j=0}^{k-1} a_j^2 \to \sigma^2\sum_{j=0}^{\infty} a_j^2 = \mathrm{E}U_t^2
$$

- 按定义可知$\{U_t\}$为纯非决定性的。
- 注意：这个证明对一般单边线性序列不适用。

Wold定理的证明—VII

- 已证明$\{V_t\}$平稳，来证明$\{V_t\}$是决定性的。用定理2.5。
- 注意$\varepsilon_{t-j} \in H_{t-j}$，

$$
\begin{aligned}
V_t =& X_t - U_t = X_t - \varepsilon_t - \sum_{j=1}^{\infty} a_j \varepsilon_{t-j} \\
=& L(X_t | H_{t-1}) - \sum_{j=1}^{\infty} a_j \varepsilon_{t-j}
\end{aligned}
$$

- 其中$L(X_t|H_{t-1}) \in H_{t-1}$，$\sum_{j=1}^{\infty} a_j \varepsilon_{t-j} \in H_\varepsilon(t-1) \subset H_{t-1}$，所以$V_t \in H_{t-1}$。

Wold定理的证明—VII

- 已证明$\{V_t\}$平稳，来证明$\{V_t\}$是决定性的。用定理2.5。
- 注意$\varepsilon_{t-j} \in H_{t-j}$，

$$
\begin{aligned}
V_t =& X_t - U_t = X_t - \varepsilon_t - \sum_{j=1}^{\infty} a_j \varepsilon_{t-j} \\
=& L(X_t | H_{t-1}) - \sum_{j=1}^{\infty} a_j \varepsilon_{t-j}
\end{aligned}
$$

- 其中$L(X_t|H_{t-1}) \in H_{t-1}$，$\sum_{j=1}^{\infty} a_j \varepsilon_{t-j} \in H_\varepsilon(t-1) \subset H_{t-1}$，所以$V_t \in H_{t-1}$。

Wold定理的证明—VII

- 已证明$\{V_t\}$平稳，来证明$\{V_t\}$是决定性的。用定理2.5。
- 注意$\varepsilon_{t-j} \in H_{t-j}$，

$$\begin{aligned} V_t =& X_t - U_t = X_t - \varepsilon_t - \sum_{j=1}^{\infty} a_j \varepsilon_{t-j} \\ =& L(X_t | H_{t-1}) - \sum_{j=1}^{\infty} a_j \varepsilon_{t-j} \end{aligned}$$

- 其中$L(X_t|H_{t-1}) \in H_{t-1}$，$\sum_{j=1}^{\infty} a_j \varepsilon_{t-j} \in H_\varepsilon(t-1) \subset H_{t-1}$，所以$V_t \in H_{t-1}$。

Wold定理的证明—VII(续)

- 注意$H_{t-1} \subset \overline{\text{sp}}\{U_s, V_s: s \le t-1\}$，由引理2.2，存在$\xi_m \in L_{\{V_{t-1},\dots,V_{t-m}\}}$，$\eta_m \in L_{\{U_{t-1},\dots,U_{t-m}\}}$，使

$$\|\xi_m + \eta_m - V_t\|^2 \to 0$$

- 已证明$\{V_t\}$与$\{U_t\}$正交所以η_m与$\xi_m - V_t$正交，于是

$$\|\xi_m + \eta_m - V_t\|^2 = \|\xi_m - V_t\|^2 + \|\eta_m\|^2 \\ \ge \|\xi_m - V_t\|^2$$

- 故$\|\xi_m - V_t\|^2 \to 0 (m \to \infty)$，由引理2.2知$V_t \in H_V(t-1) = \overline{\text{sp}}\{V_s: s \le t-1\}$。
- 由定理2.5知$\{V_t\}$为决定性的，且$H_V(t) = H_V(t-j), j \in \mathbb{Z}$，所以$V_t \in H_V(t-j)$，$t, j \in \mathbb{Z}$。

Wold定理的证明—VII(续)

- 注意$H_{t-1} \subset \overline{\text{sp}}\{U_s, V_s: \ s \le t-1\}$，由引理2.2，存在$\xi_m \in L_{\{V_{t-1},\ldots,V_{t-m}\}}$，$\eta_m \in L_{\{U_{t-1},\ldots,U_{t-m}\}}$，使

$$\|\xi_m + \eta_m - V_t\|^2 \to 0$$

- 已证明$\{V_t\}$与$\{U_t\}$正交所以η_m与$\xi_m - V_t$正交，于是

$$\begin{aligned}\|\xi_m + \eta_m - V_t\|^2 &= \|\xi_m - V_t\|^2 + \|\eta_m\|^2 \\ \ge &\|\xi_m - V_t\|^2\end{aligned}$$

- 故$\|\xi_m - V_t\|^2 \to 0 (m \to \infty)$，由引理2.2知$V_t \in H_V(t-1) = \overline{\text{sp}}\{V_s: \ s \le t-1\}$。
- 由定理2.5知$\{V_t\}$为决定性的，且$H_V(t) = H_V(t-j), j \in \mathbb{Z}$，所以$V_t \in H_V(t-j)$，$t, j \in \mathbb{Z}$。

Wold定理的证明—VII(续)

- 注意$H_{t-1} \subset \overline{\text{sp}}\{U_s, V_s: \ s \leq t-1\}$，由引理2.2，存在$\xi_m \in L_{\{V_{t-1},\dots,V_{t-m}\}}$，$\eta_m \in L_{\{U_{t-1},\dots,U_{t-m}\}}$，使

$$\|\xi_m + \eta_m - V_t\|^2 \to 0$$

- 已证明$\{V_t\}$与$\{U_t\}$正交所以η_m与$\xi_m - V_t$正交，于是

$$\begin{aligned}&\|\xi_m + \eta_m - V_t\|^2 = \|\xi_m - V_t\|^2 + \|\eta_m\|^2\\ \geq&\|\xi_m - V_t\|^2\end{aligned}$$

- 故$\|\xi_m - V_t\|^2 \to 0 (m \to \infty)$，由引理2.2知$V_t \in H_V(t-1) = \overline{\text{sp}}\{V_s: \ s \leq t-1\}$。
- 由定理2.5知$\{V_t\}$为决定性的，且$H_V(t) = H_V(t-j), j \in \mathbb{Z}$，所以$V_t \in H_V(t-j)$，$t, j \in \mathbb{Z}$。

Wold定理的证明—VII(续)

- 注意$H_{t-1} \subset \overline{\text{sp}}\{U_s, V_s: \ s \leq t-1\}$，由引理2.2，存在$\xi_m \in L_{\{V_{t-1},\dots,V_{t-m}\}}$，$\eta_m \in L_{\{U_{t-1},\dots,U_{t-m}\}}$，使

$$\|\xi_m + \eta_m - V_t\|^2 \to 0$$

- 已证明$\{V_t\}$与$\{U_t\}$正交所以η_m与$\xi_m - V_t$正交，于是

$$\begin{aligned} &\|\xi_m + \eta_m - V_t\|^2 = \|\xi_m - V_t\|^2 + \|\eta_m\|^2 \\ \geq &\|\xi_m - V_t\|^2 \end{aligned}$$

- 故$\|\xi_m - V_t\|^2 \to 0 (m \to \infty)$，由引理2.2知$V_t \in H_V(t-1) = \overline{\text{sp}}\{V_s: \ s \leq t-1\}$。
- 由定理2.5知$\{V_t\}$为决定性的，且$H_V(t) = H_V(t-j), j \in \mathbb{Z}$，所以$V_t \in H_V(t-j)$，$t, j \in \mathbb{Z}$。

关于新息的讨论

- $\varepsilon_t = X_t - L(X_t|H_{t-1})$是$X_t$提供的比$X_{t-1}, X_{t-2}, \ldots$多的信息(线性意义下)。
- 可以证明

$$H_t = \overline{\mathrm{sp}}(\varepsilon_t) \oplus H_{t-1} \qquad (*)$$

- 这样

$$H_t = \bigoplus_{j=0}^{\infty} \overline{\mathrm{sp}}(\varepsilon_{t-j}) \oplus H_{-\infty}$$

其中$H_{-\infty} = \bigcap_t H_t$。$V_t \in H_{-\infty}$。

关于新息的讨论

- $\varepsilon_t = X_t - L(X_t|H_{t-1})$是$X_t$提供的比$X_{t-1}, X_{t-2}, \ldots$多的信息(线性意义下)。
- 可以证明

$$H_t = \overline{\text{sp}}(\varepsilon_t) \oplus H_{t-1} \qquad (*)$$

- 这样

$$H_t = \bigoplus_{j=0}^{\infty} \overline{\text{sp}}(\varepsilon_{t-j}) \oplus H_{-\infty}$$

其中$H_{-\infty} = \bigcap_t H_t$。$V_t \in H_{-\infty}$。

关于新息的讨论

- $\varepsilon_t = X_t - L(X_t|H_{t-1})$是$X_t$提供的比$X_{t-1}, X_{t-2}, \ldots$多的信息(线性意义下)。
- 可以证明

$$H_t = \overline{\text{sp}}(\varepsilon_t) \oplus H_{t-1} \qquad (*)$$

- 这样

$$H_t = \bigoplus_{j=0}^{\infty} \overline{\text{sp}}(\varepsilon_{t-j}) \oplus H_{-\infty}$$

其中$H_{-\infty} = \bigcap_t H_t$。$V_t \in H_{-\infty}$。

关于新息的讨论(续)

- 事实上，(*)右侧两项正交且都属于H_t，记$\tilde{H}_t = \overline{\text{sp}}(\{\varepsilon_t\} \cup H_{t-1})$，则$\tilde{H}_t \subset H_t$，来证$H_t \subset \tilde{H}_t$。
- 由引理2.2只要证明$X_s \in \tilde{H}_t, s \leq t$。当$s < t$时显然，对$X_t$，因为

$$X_t = \varepsilon_t + L(X_t | H_{t-1})$$

所以X_t是$\{\varepsilon_t\} \cup H_{t-1}$中元素的线性组合。
- 因此

$$H_t = \tilde{H}_t = \overline{\text{sp}}(\{\varepsilon_t\} \cup H_{t-1})$$

- 再由$\varepsilon_t \perp H_{t-1}$可得$H_t = \overline{\text{sp}}(\varepsilon_t) \oplus H_{t-1}$。

关于新息的讨论(续)

- 事实上，(*)右侧两项正交且都属于H_t，记$\tilde{H}_t = \overline{\text{sp}}(\{\varepsilon_t\} \cup H_{t-1})$, 则$\tilde{H}_t \subset H_t$，来证$H_t \subset \tilde{H}_t$。
- 由引理2.2只要证明$X_s \in \tilde{H}_t, s \le t$。当$s < t$时显然，对$X_t$，因为

$$X_t = \varepsilon_t + L(X_t | H_{t-1})$$

所以X_t是$\{\varepsilon_t\} \cup H_{t-1}$中元素的线性组合。
- 因此

$$H_t = \tilde{H}_t = \overline{\text{sp}}(\{\varepsilon_t\} \cup H_{t-1})$$

- 再由$\varepsilon_t \perp H_{t-1}$可得$H_t = \overline{\text{sp}}(\varepsilon_t) \oplus H_{t-1}$。

关于新息的讨论(续)

- 事实上，(*)右侧两项正交且都属于H_t，记$\tilde{H}_t = \overline{\text{sp}}(\{\varepsilon_t\} \cup H_{t-1})$，则$\tilde{H}_t \subset H_t$，来证$H_t \subset \tilde{H}_t$。
- 由引理2.2只要证明$X_s \in \tilde{H}_t, s \le t$。当$s < t$时显然，对$X_t$，因为

$$X_t = \varepsilon_t + L(X_t | H_{t-1})$$

所以X_t是$\{\varepsilon_t\} \cup H_{t-1}$中元素的线性组合。
- 因此

$$H_t = \tilde{H}_t = \overline{\text{sp}}(\{\varepsilon_t\} \cup H_{t-1})$$

- 再由$\varepsilon_t \perp H_{t-1}$可得$H_t = \overline{\text{sp}}(\varepsilon_t) \oplus H_{t-1}$。

关于新息的讨论(续)

- 事实上，(*)右侧两项正交且都属于H_t，
  记$\tilde{H}_t = \overline{\text{sp}}(\{\varepsilon_t\} \cup H_{t-1})$，则$\tilde{H}_t \subset H_t$，来证$H_t \subset \tilde{H}_t$。
- 由引理2.2只要证明$X_s \in \tilde{H}_t, s \leq t$。当$s < t$时显然，对$X_t$，因为

$$X_t = \varepsilon_t + L(X_t | H_{t-1})$$

  所以X_t是$\{\varepsilon_t\} \cup H_{t-1}$中元素的线性组合。
- 因此

$$H_t = \tilde{H}_t = \overline{\text{sp}}(\{\varepsilon_t\} \cup H_{t-1})$$

- 再由$\varepsilon_t \perp H_{t-1}$可得$H_t = \overline{\text{sp}}(\varepsilon_t) \oplus H_{t-1}$。

多步预报的均方误差

- 设$\{X_t\}$是非决定性的平稳列，由Wold表示定理(5)，$V_{t+n} \in H_t$，所以用无穷长历史进行的最佳线性预测为

$$\begin{aligned} L(X_{t+n}|H_t) =& L(U_{t+n}|H_t) + L(V_{t+n}|H_t) \\ =& L(\sum_{j=0}^{\infty} a_j \varepsilon_{t+n-j}|H_t) + V_{t+n} \\ =& \sum_{j=n}^{\infty} a_j \varepsilon_{t+n-j} + V_{t+n} \end{aligned} \tag{2.20}$$

- 称$L(X_{t+n}|H_t)$是X_{t+n}的n步预报，由Wold分解公式知预报误差为

$$X_{t+n} - L(X_{t+n}|H_t) = \sum_{j=0}^{n-1} a_j \varepsilon_{t+n-j} \tag{2.21}$$

- 预报的均方误差为

$$\sigma^2(n) = \sigma^2 \sum_{j=0}^{n-1} a_j^2 \tag{2.22}$$

- $n \to \infty$时$\sigma^2(n) \to \mathrm{E}U_t^2$.

多步预报的均方误差

- 设$\{X_t\}$是非决定性的平稳列，由Wold表示定理(5)，$V_{t+n} \in H_t$，所以用无穷长历史进行的最佳线性预测为

$$\begin{aligned} L(X_{t+n}|H_t) =& L(U_{t+n}|H_t) + L(V_{t+n}|H_t) \\ =& L(\sum_{j=0}^{\infty} a_j \varepsilon_{t+n-j}|H_t) + V_{t+n} \\ =& \sum_{j=n}^{\infty} a_j \varepsilon_{t+n-j} + V_{t+n} \end{aligned} \tag{2.20}$$

- 称$L(X_{t+n}|H_t)$是X_{t+n}的n步预报，由Wold分解公式知预报误差为

$$X_{t+n} - L(X_{t+n}|H_t) = \sum_{j=0}^{n-1} a_j \varepsilon_{t+n-j} \tag{2.21}$$

- 预报的均方误差为

$$\sigma^2(n) = \sigma^2 \sum_{j=0}^{n-1} a_j^2 \tag{2.22}$$

- $n \to \infty$时$\sigma^2(n) \to \mathrm{E}U_t^2$.

多步预报的均方误差

- 设$\{X_t\}$是非决定性的平稳列，由Wold表示定理(5)，$V_{t+n} \in H_t$，所以用无穷长历史进行的最佳线性预测为

$$\begin{aligned} L(X_{t+n}|H_t) =& L(U_{t+n}|H_t) + L(V_{t+n}|H_t) \\ =& L(\sum_{j=0}^{\infty} a_j \varepsilon_{t+n-j}|H_t) + V_{t+n} \\ =& \sum_{j=n}^{\infty} a_j \varepsilon_{t+n-j} + V_{t+n} \end{aligned} \tag{2.20}$$

- 称$L(X_{t+n}|H_t)$是X_{t+n}的n步预报，由Wold分解公式知预报误差为

$$X_{t+n} - L(X_{t+n}|H_t) = \sum_{j=0}^{n-1} a_j \varepsilon_{t+n-j} \tag{2.21}$$

- 预报的均方误差为

$$\sigma^2(n) = \sigma^2 \sum_{j=0}^{n-1} a_j^2 \tag{2.22}$$

- $n \to \infty$时$\sigma^2(n) \to \mathrm{E}U_t^2$.

多步预报的均方误差

- 设$\{X_t\}$是非决定性的平稳列，由Wold表示定理(5)，$V_{t+n} \in H_t$，所以用无穷长历史进行的最佳线性预测为

$$\begin{aligned} L(X_{t+n}|H_t) =& L(U_{t+n}|H_t) + L(V_{t+n}|H_t) \\ =& L(\sum_{j=0}^{\infty} a_j \varepsilon_{t+n-j}|H_t) + V_{t+n} \\ =& \sum_{j=n}^{\infty} a_j \varepsilon_{t+n-j} + V_{t+n} \end{aligned} \tag{2.20}$$

- 称$L(X_{t+n}|H_t)$是X_{t+n}的n步预报，由Wold分解公式知预报误差为

$$X_{t+n} - L(X_{t+n}|H_t) = \sum_{j=0}^{n-1} a_j \varepsilon_{t+n-j} \tag{2.21}$$

- 预报的均方误差为

$$\sigma^2(n) = \sigma^2 \sum_{j=0}^{n-1} a_j^2 \tag{2.22}$$

- $n \to \infty$时$\sigma^2(n) \to \mathrm{E}U_t^2$.

Kolmogorov公式

- **定理2.7** (Kolmogorov公式[7]) 设$\{U_t\}$是非决定性平稳序列$\{X_t\}$的纯非决定性部分，$f(\lambda)$是$\{U_t\}$的谱密度. 则有

$$\sigma^2 = E[X_t - L(X_t|H_{t-1})]^2 = 2\pi \exp\left(\frac{1}{2\pi}\int_{-\pi}^{\pi} \ln f(\lambda) d\lambda\right). \quad (2.24)$$

- 公式(2.24)的证明需要较多解析函数的知识.
- 当$\{U_t\}$是白噪声时，公式(2.24)明显是成立的.
- 从Kolmogorov公式(2.24)看到，如果$\{X_t\}$是非决定性的，则它的纯非决定性部分的谱密度$f(\lambda)$必是ln可积的，即

$$\int_{-\pi}^{\pi} \ln f(\lambda) d\lambda > -\infty. \quad (2.25)$$

Kolmogorov公式

- **定理2.7** (Kolmogorov公式[7]) 设$\{U_t\}$是非决定性平稳序列$\{X_t\}$的纯非决定性部分，$f(\lambda)$是$\{U_t\}$的谱密度. 则有

$$\sigma^2 = E[X_t - L(X_t|H_{t-1})]^2 = 2\pi \exp\left(\frac{1}{2\pi}\int_{-\pi}^{\pi} \ln f(\lambda) d\lambda\right). \tag{2.24}$$

- 公式(2.24)的证明需要较多解析函数的知识.
- 当$\{U_t\}$是白噪声时，公式(2.24)明显是成立的.
- 从Kolmogorov公式(2.24)看到，如果$\{X_t\}$是非决定性的，则它的纯非决定性部分的谱密度$f(\lambda)$必是ln可积的，即

$$\int_{-\pi}^{\pi} \ln f(\lambda) d\lambda > -\infty. \tag{2.25}$$

Kolmogorov公式

- **定理2.7** (Kolmogorov公式[7]) 设$\{U_t\}$是非决定性平稳序列$\{X_t\}$的纯非决定性部分，$f(\lambda)$是$\{U_t\}$的谱密度. 则有

$$\sigma^2 = E[X_t - L(X_t|H_{t-1})]^2 = 2\pi \exp\left(\frac{1}{2\pi}\int_{-\pi}^{\pi} \ln f(\lambda) d\lambda\right). \tag{2.24}$$

- 公式(2.24)的证明需要较多解析函数的知识.
- 当$\{U_t\}$是白噪声时，公式(2.24)明显是成立的.
- 从Kolmogorov公式(2.24)看到，如果$\{X_t\}$是非决定性的，则它的纯非决定性部分的谱密度$f(\lambda)$必是ln可积的，即

$$\int_{-\pi}^{\pi} \ln f(\lambda) d\lambda > -\infty. \tag{2.25}$$

Kolmogorov公式

- **定理2.7** (Kolmogorov公式[7]) 设$\{U_t\}$是非决定性平稳序列$\{X_t\}$的纯非决定性部分，$f(\lambda)$是$\{U_t\}$的谱密度. 则有

$$\sigma^2 = E[X_t - L(X_t|H_{t-1})]^2 = 2\pi \exp\left(\frac{1}{2\pi}\int_{-\pi}^{\pi} \ln f(\lambda) d\lambda\right). \tag{2.24}$$

- 公式(2.24)的证明需要较多解析函数的知识.
- 当$\{U_t\}$是白噪声时，公式(2.24)明显是成立的.
- 从Kolmogorov公式(2.24)看到，如果$\{X_t\}$是非决定性的，则它的纯非决定性部分的谱密度$f(\lambda)$必是ln可积的，即

$$\int_{-\pi}^{\pi} \ln f(\lambda) d\lambda > -\infty. \tag{2.25}$$

最佳预测

- 设$\{X_t\}$是平稳序列, 用$\mathscr{F}_t = \sigma\{X_t, X_{t-1}, \dots\}$ 表示由$X_t, X_{t-1}, \dots$生成的σ-代数. 称条件数学期望

$$E(X_{t+k}|\mathscr{F}_t)$$

是用全体历史$\{X_j : j \le t\}$对X_{t+k}进行预测时的最佳预测.

- 最佳预测是均方误差最小的，这是因为条件数学期望$E(X_{t+k}|\mathscr{F}_t)$是$X_{t-1}, X_{t-2}, \dots$的函数, 二阶矩有限:

$$E[E(X_{t+k}|\mathscr{F}_t)]^2 \le E[E(X_{t+k}^2|\mathscr{F}_t)] = EX_{t+k}^2 < \infty.$$

由概率论中数学期望性质可以证明对任意二阶矩有限的$X_t, X_{t-1}, \dots$的函数ξ有

$$\mathrm{E}(X_{t+k} - \xi)^2 \ge \mathrm{E}[X_{t+k} - \mathrm{E}(X_{t+k}|\mathscr{F}_t)]^2 \qquad (2.26)$$

最佳预测

- 设$\{X_t\}$是平稳序列，用$\mathscr{F}_t = \sigma\{X_t, X_{t-1}, \dots\}$ 表示由$X_t, X_{t-1}, \dots$生成的σ-代数. 称条件数学期望

$$E(X_{t+k}|\mathscr{F}_t)$$

是用全体历史$\{X_j : j \le t\}$对X_{t+k}进行预测时的最佳预测.

- 最佳预测是均方误差最小的，这是因为条件数学期望$E(X_{t+k}|\mathscr{F}_t)$是$X_{t-1}, X_{t-2}, \dots$的函数，二阶矩有限:

$$E[E(X_{t+k}|\mathscr{F}_t)]^2 \le E[E(X_{t+k}^2|\mathscr{F}_t)] = EX_{t+k}^2 < \infty.$$

由概率论中数学期望性质可以证明对任意二阶矩有限的$X_t, X_{t-1}, \dots$的函数ξ有

$$\mathrm{E}(X_{t+k} - \xi)^2 \ge \mathrm{E}[X_{t+k} - \mathrm{E}(X_{t+k}|\mathscr{F}_t)]^2 \qquad (2.26)$$

最佳预测和最佳线性预测相等的条件

- 最佳预测一般比最佳线性预测好，但是对纯非决定性序列如果其新息是独立序列则二者等价。
- **定理2.8**　设平稳序列$\{X_t\}$有Wold表示

$$X_t = \sum_{j=0}^{\infty} a_j \varepsilon_{t-j}, \quad t \in \mathbb{Z}. \tag{2.27}$$

则

$$L(X_{t+n}|H_t) = E(X_{t+n}|\mathscr{F}_t),\ n \geq 1, \quad t \in \mathbb{Z}, \tag{2.28}$$

成立的充分必要条件是

$$E(\varepsilon_{t+1}|\varepsilon_t, \varepsilon_{t-1}, \ldots) = 0, \quad t \in \mathbb{Z}. \tag{2.29}$$

- (2.29)的条件称为鞅差。

最佳预测和最佳线性预测相等的条件

- 最佳预测一般比最佳线性预测好，但是对纯非决定性序列如果其新息是独立序列则二者等价。
- **定理2.8** 设平稳序列$\{X_t\}$有Wold表示

$$X_t = \sum_{j=0}^{\infty} a_j \varepsilon_{t-j}, \quad t \in \mathbb{Z}. \tag{2.27}$$

则

$$L(X_{t+n}|H_t) = E(X_{t+n}|\mathscr{F}_t), \ n \geq 1, \quad t \in \mathbb{Z}, \tag{2.28}$$

成立的充分必要条件是

$$E(\varepsilon_{t+1}|\varepsilon_t, \varepsilon_{t-1}, \ldots) = 0, \quad t \in \mathbb{Z}. \tag{2.29}$$

- (2.29)的条件称为鞅差。

最佳预测和最佳线性预测相等的条件

- 最佳预测一般比最佳线性预测好，但是对纯非决定性序列如果其新息是独立序列则二者等价。
- **定理2.8** 设平稳序列$\{X_t\}$有Wold表示

$$X_t = \sum_{j=0}^{\infty} a_j \varepsilon_{t-j}, \quad t \in \mathbb{Z}. \tag{2.27}$$

则

$$L(X_{t+n}|H_t) = E(X_{t+n}|\mathscr{F}_t), \; n \geq 1, \quad t \in \mathbb{Z}, \tag{2.28}$$

成立的充分必要条件是

$$E(\varepsilon_{t+1}|\varepsilon_t, \varepsilon_{t-1}, \dots) = 0, \quad t \in \mathbb{Z}. \tag{2.29}$$

- (2.29)的条件称为鞅差。

关于ARMA模型的推论

- **推论2.9** 设ARMA(p,q)序列$\{X_t\}$中的新息$\{\varepsilon_t\}$是独立白噪声，则用全体历史$\{X_t, X_{t-1}, \ldots\}$对X_{t+n}进行预测时，最佳预测和最佳线性预测相等.