# Chapter 2 Lecture in STAT 730, Spring 2026

Here: Stochastic Regression ideas for Shumway-Stoffer Chap. 2.

Previous lecture: stochastic regression via Partial Likelihood as in Slud & Kedem (1994, *Stat. Sinica*) and Kedem & Fokianos, **Regression Models in Time Series Analysis** (2202), especially conditionally specified Logistic Regression (+ other categorical time series models) with time-dependent covariates.

Topics in Ch. 2: classic linear regression models,
search for trends and periodicities.

## Linear Regression as One-Step-Ahead Model

Model is $\begin{pmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_n \end{pmatrix} = \mathbf{Z}\,\beta + \underline{\epsilon}\ , \quad \mathbf{Z} = \Big( \underset{n\times p}{\mathbf{X}} \,\Big|\, \underset{n\times q}{\mathbf{X}^*} \Big)\ , \ \ \beta = (\gamma, \gamma^*)$

**Slides PLregrTalk2023.pdf** introduce **Partial Likelihood** idea.

OLS Regression: (treats $\epsilon_t \overset{iid}{\sim} (0, \sigma^2)$ even if it isn't)

$$\text{logPL} = \prod_{t=1}^{n} f_{Y_t \,|\, (Y_s, Z_s,\, s<t)}(\beta, \sigma^2) = -\frac{n}{2}\log(\sigma^2) - \frac{1}{2\sigma^2}\|\mathbf{Y} - \mathbf{Z}\beta\|^2$$

Assume that $X_t$, $X_t^*$ processes are *adapted* to the timing of $Y_t$ in the sense that $(X_s, X_s^*, \ s \le t)$ are (determined by $\mathcal{F}_{t-1}$) observable up through time $t-1$

# OLS Estimation and Asymptotic Theory

Model: $\epsilon_t = Y_t - \beta^{tr} \mathbf{Z}_t$ , $E(\epsilon_t \,|\, \mathcal{F}_{t-1}) = 0$, $E(\epsilon_t^2 \,|\, \mathcal{F}_{t-1}) = \sigma^2$

$$\sum_{s=1}^{t} \mathbf{Z}_t \, \epsilon_t \quad \text{is a martingale} \qquad \textit{recall def'n}$$

Conditional Estimating Equation $\quad \mathbf{Z}^{tr} (\mathbf{Y} - \mathbf{Z}\widehat{\beta}) = \underline{0}$

$$SSE = \|\mathbf{Y} - \mathbf{Z}\widehat{\beta}\|^2 = \mathbf{Y}^{tr}(\mathbf{I} - \mathbf{Z}(\mathbf{Z}^{tr}\mathbf{Z})^{-1}\mathbf{Z})\mathbf{Y}$$

$$\widehat{\beta} - \beta = (\mathbf{Z}^{tr}\mathbf{Z})^{-1} \mathbf{Z}^{tr} \epsilon$$

In stochastic-regression, RHS not 'causal' (coef. of $\epsilon_t$ in $\mathcal{F}_{t-1}$)

but $\frac{1}{n} \mathbf{Z}^{tr}\mathbf{Z} = \frac{1}{n} \sum_{t=1}^{n} (Z_t)^{\otimes 2}$ may $\xrightarrow{P} V$, e.g., to $E(Z_1^{\otimes 2})$

Under conditions: $\quad \frac{1}{\sqrt{n}}(\widehat{\beta} - \beta) \xrightarrow{\mathcal{D}} \mathcal{N}_{p+q}(\underline{0}, \, \sigma^2 \, V^{-1})$

## ANOVA-style Model Comparisons

Consider null hypothesis $H_0 : \gamma^* = \underline{0}$ that $\mathbf{X}^*$ can be dropped

Analyze under $H_0$: $SSE^{(r)} = \|\mathbf{Y} - \mathbf{X}\widehat{\gamma}^{(r)}\|^2$

($^{(r)}$ stands for 'restricted' OLS under $H_0$)

and then consider incremental MSE $= \left(\|\mathbf{Y} - \mathbf{Z}\widehat{\beta}\|^2 - SSE^{(r)}\right) / (q\widehat{\sigma}^{(r)2})$

leads to ideas of penalized logPL optimization via AIC, BIC

# Models with Deterministic Trend

**Example 1:** polynomial trends $\mathbf{X} = \left(\mathbf{1} \,\middle|\, \{t\}_{t=1}^{n}\right)$, $\mathbf{X}^* = \{t^2\}_{t=1}^{n}$

Could formulate as regression, test if $\mathbf{X}^*$ is necessary, **or**

difference once or twice: $(I - B)Y_t \equiv Y_t - Y_{t-1}$, or

$(I - B)^2 Y_t \equiv Y_t - 2Y_{t-1} + Y_{t-2}$ and

analyze as stationary process under different assumptions on $\epsilon_t$

## Models Tailored to (Sin-Cosine) Bases

**Example 2:** sinusoidal trends, regress on

$$\mathbf{X} = \left( \{\cos(2\pi t\lambda_0)\}_{t=1}^{n} \,\Big|\, \{\sin(2\pi t\lambda_0)\}_{t=1}^{n} \right)$$

to get coefficients $\widehat{\beta}_1(\lambda_0)$, $\widehat{\beta}_2(\lambda_0)$.

*Frequency $\lambda_0$ in place of period $\omega_0 = 1/\lambda_0$*

When $\lambda_0 = j/n$, array of resulting coefficients $\widehat{\beta}_1(j,n)$, $\widehat{\beta}_2(j,n)$ are closely related to the **periodogram** $\widehat{\beta}_1^2(j/n)+\widehat{\beta}_2^2(j/n)$ plotted against $j/n$ as in Fig. 2.10.