Lecture 28

Dynamic Linear Model: Part 1

Arnab Hazra

Binary responses (non-time series): Logistic and probit regressions

- The response variable Research can take values only zero or one. Suppose, they are denoted by $Z_i, i = 1, \ldots, 400$. Also, suppose, for the new student, it is denoted by Z_0.

	GRE	TOEFL	Research
1	337	118	1
2	324	107	1
3	316	104	1
4	322	110	1
5	314	103	0
6	330	115	1

...

Binary responses (non-time series): Logistic and probit regressions

- The predicted response should also be either zero or one. Other values are meaningless.

- Instead of the value of the response, it is rather more meaningful to quantify $\Pr(Z_0 = 0|X_0 = 110)$ or $\Pr(Z_0 = 1|X_0 = 110)$.

- Thus, the most reasonable approach for quantifying the linear relationship between Z_i's and X_i's would be to choose

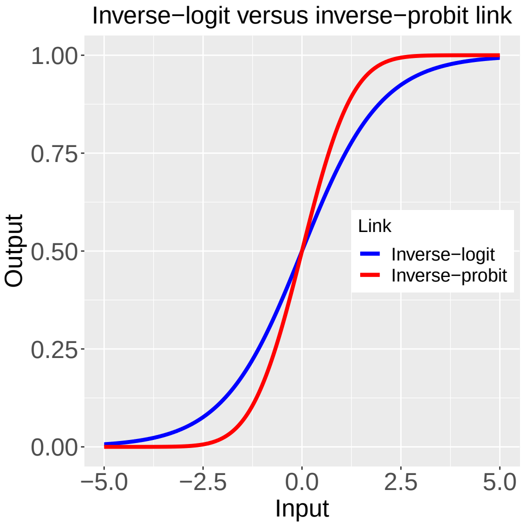
$$\Pr(Z_0 = 1|X_0 = 110) = \eta(\beta_0 + \beta_1 X_i),$$

where $\eta : \mathbb{R} \to [0, 1]$ and monotone.

Binary responses (non-time series): Logistic and probit regressions

- We call the regression model to be logistic if $\eta(\cdot)$ is a inverse-logit link, and probit if $\eta(\cdot)$ is a inverse-probit link. The functions are defined as follows.
    - Logistic regression: $\eta(x) = \exp[x]/(1 + \exp[x])$
    - Probit regression: $\eta(x) = \Phi(x)$, where $\Phi(\cdot)$ denotes the standard normal CDF

Binary responses (non-time series): Logistic and probit regressions



Hierarchical structure of probit regressions

- Suppose, $W|X = x \sim \mathrm{Normal}(\beta_0 + \beta_1 x, 1)$, and $Y = 1$ if $W > 0$, $Y = 0$ if $W < 0$. Show that there exists a inverse-probit link between X and Y, that is, $\Pr(Y = 1|X = x) = \Phi(\beta_0 + \beta_1 x)$.

$$
\begin{aligned}
& \Pr(Y = 1|X = x) \\
= \; & \Pr(W > 0|X = x) \\
= \; & 1 - \Pr(W \leq 0|X = x) \\
= \; & 1 - \Pr(W - \beta_0 - \beta_1 x \leq -\beta_0 - \beta_1 x|X = x) \\
= \; & 1 - \Phi(-\beta_0 - \beta_1 x) \\
= \; & \Phi(\beta_0 + \beta_1 x)
\end{aligned}
$$

Motivation of hierarchical models

- Now suppose we have a time series of binary observations $Y_t \in \{0,1\}$.
- It is difficult to define a joint distribution of $(Y_1, \ldots, Y_T)'$.
- Suppose $X_1, \ldots, X_T$ are covariate values.
- However we have seen

$$
\begin{aligned}
& \pi(Y_1, \ldots, Y_T) \\
= & \int \cdots \int \pi(Y_1, \ldots, Y_T, W_1, \ldots, W_T) dW_1 \ldots dW_T \\
= & \int \cdots \int \pi(Y_1, \ldots, Y_T | W_1, \ldots, W_T) \pi(W_1, \ldots, W_T) dW_1 \ldots dW_T \\
= & \int \cdots \int \left\{ \prod_{t=1}^{T} \pi(Y_t | W_t) \right\} \pi(W_1, \ldots, W_T) dW_1 \ldots dW_T \\
\neq & \prod_{t=1}^{T} \pi(Y_t)
\end{aligned}
$$

Motivation of hierarchical models (contd.)

- Suppose, $W_t|X_t = x_t \sim \mathrm{Normal}(\beta_0 + \beta_1 x_t, 1)$, and $Y_t = 1$ if $W_t > 0$, $Y_t = 0$ if $W_t < 0$. Then $\Pr(Y_t = 1|X_t = x_t) = \Phi(\beta_0 + \beta_1 x_t)$.

- Now we can choose a time series model for W_t. For example, we can choose an AR(1) model for W_t.

- In that case, we define $W_t^* = W_t - \beta_0 + \beta_1 x_t$, and then $W_1 \sim \mathcal{N}(0, 1)$ and $W_t^*|W_{t-1}^* \sim \mathcal{N}(\rho W_{t-1}^*, 1 - \rho^2)$. Note that $W_t^* \sim \mathcal{N}(0, 1)$ for all t.

- We can choose a more general model like ARMA(p, q) but the marginal needs to be standard normal.

DLM: State equation

- DLM in its basic form, employs an order one, p-dimensional vector autoregression as the state equation,

$$\boldsymbol{X}_t = \Phi \boldsymbol{X}_{t-1} + \boldsymbol{W}_t,$$

where $\boldsymbol{W}_t \overset{IID}{\sim} \mathcal{N}_p(\boldsymbol{0}, \boldsymbol{Q})$.

- In the DLM, we assume the process starts with a normal vector $\boldsymbol{X}_0$, such that $\boldsymbol{X}_0 \sim \mathcal{N}_p(\boldsymbol{\mu}_0, \Sigma_0)$.

- Here p is called state dimension.

DLM: Observation equation

- We do not observe the state vector $\boldsymbol{X}_t$ directly, but only a linear transformed version of it with noise added, say $\boldsymbol{Y}_t = \boldsymbol{A}_t \boldsymbol{X}_t + \boldsymbol{V}_t$, where $\boldsymbol{A}_t$ is a $q \times p$ measurement or observation matrix; this equation is called the observation equation.

- The observed data vector, $\boldsymbol{Y}_t$, is q-dimensional, which can be larger than or smaller than p, the state dimension. The additive observation noise is $\boldsymbol{V}_t \overset{IID}{\sim} \mathcal{N}_q(\mathbf{0}, \boldsymbol{R})$.

Thank you!