Linear Predictors

Econ 211C – Unit 2, Section 1

Eric M. Aldrich

UC Santa Cruz

Forecasting

Suppose we are interested in forecasting a random variable Y_{t+1} based on a set of variables $\boldsymbol{X}_t$.

- $\boldsymbol{X}_t$ might be comprised of m lags of Y_{t+1}: $Y_t, Y_{t-1}, \ldots, Y_{t-m+1}$.
- We can denote $Y^*_{t+1|t}$ as the forecast of Y_{t+1} based on $\boldsymbol{X}_t$.
- We can choose $Y^*_{t+1|t}$ to minimize some loss function, $L\left(Y^*_{t+1|t}\right)$, which evaluates the quality of $Y^*_{t+1|t}$.
- A common choice is the quadratic loss function:

$$L\left(Y^*_{t+1|t}\right) = \mathrm{E}\left[\left(Y_{t+1} - Y^*_{t+1|t}\right)^2\right].$$

Mean Squared Error Loss

Quadratic loss is also known as *mean squared error*.

$$MSE\left(Y^*_{t+1|t}\right) = \mathrm{E}\left[\left(Y_{t+1} - Y^*_{t+1|t}\right)^2\right].$$

- The conditional expectation, $\mathrm{E}\left[Y_{t+1}|\boldsymbol{X}_t\right]$ minimizes $MSE\left(Y^*_{t+1|t}\right)$.

MSE Minimizer

Let $Y^*_{t+1|t} = g(\boldsymbol{X}_t)$. Then

$$
\begin{aligned}
\mathrm{E}\left[\left(Y_{t+1} - g(\boldsymbol{X}_t)\right)^2\right] &= \mathrm{E}\Big[\big(Y_{t+1} - \mathrm{E}\left[Y_{t+1}|\boldsymbol{X}_t\right] \\
&\qquad\qquad + \mathrm{E}\left[Y_{t+1}|\boldsymbol{X}_t\right] - g(\boldsymbol{X}_t)\big)^2\Big] \\
&= \mathrm{E}\left[\left(Y_{t+1} - \mathrm{E}\left[Y_{t+1}|\boldsymbol{X}_t\right]\right)^2\right] \\
&\quad + 2\mathrm{E}\Big[\big(Y_{t+1} - \mathrm{E}\left[Y_{t+1}|\boldsymbol{X}_t\right]\big) \\
&\qquad\qquad \times \big(\mathrm{E}\left[Y_{t+1}|\boldsymbol{X}_t\right] - g(\boldsymbol{X}_t)\big)\Big] \\
&\qquad\qquad + \mathrm{E}\left[\big(\mathrm{E}\left[Y_{t+1}|\boldsymbol{X}_t\right] - g(\boldsymbol{X}_t)\big)^2\right]
\end{aligned}
$$

MSE Minimizer

By the law of iterated expectations

$$\begin{aligned}
&\mathrm{E}\left[\left(Y_{t+1} - \mathrm{E}\left[Y_{t+1}|\boldsymbol{X}_t\right]\right)\left(\mathrm{E}\left[Y_{t+1}|\boldsymbol{X}_t\right] - g(\boldsymbol{X}_t)\right)\right] \\
&\qquad = \mathrm{E}\left[\mathrm{E}\left[\,(Y_{t+1} - \mathrm{E}\left[Y_{t+1}|\boldsymbol{X}_t\right])\,\middle|\boldsymbol{X}_t\right]\left(\mathrm{E}\left[Y_{t+1}|\boldsymbol{X}_t\right] - g(\boldsymbol{X}_t)\right)\right] \\
&\qquad = \mathrm{E}\left[\,(\mathrm{E}\left[Y_{t+1}|\boldsymbol{X}_t\right] - \mathrm{E}\left[Y_{t+1}|\boldsymbol{X}_t\right])\left(\mathrm{E}\left[Y_{t+1}|\boldsymbol{X}_t\right] - g(\boldsymbol{X}_t)\right)\right] \\
&\qquad = 0.
\end{aligned}$$

- This means that the second term of the equation on the previous slide is zero.

MSE Minimizer

Substituting the previous result:

$$\mathrm{E}\left[(Y_{t+1} - g(\boldsymbol{X}_t))^2\right] = \mathrm{E}\left[(Y_{t+1} - \mathrm{E}\left[Y_{t+1}|\boldsymbol{X}_t\right])^2\right] + \mathrm{E}\left[\left(\mathrm{E}\left[Y_{t+1}|\boldsymbol{X}_t\right] - g(\boldsymbol{X}_t)\right)^2\right]$$

- Clearly the the MSE is minimized when

$$\mathrm{E}\left[\left(\mathrm{E}\left[Y_{t+1}|\boldsymbol{X}_t\right] - g(\boldsymbol{X}_t)\right)^2\right] = 0.$$

- This occurs when $\mathrm{E}\left[Y_{t+1}|\boldsymbol{X}_t\right] = g(\boldsymbol{X}_t)$.

Linear Projection

We can restrict our forecast to be a linear function of $\boldsymbol{X}_t$:

$$Y^*_{t+1|t} = \boldsymbol{X}_t'\boldsymbol{\beta}.$$

- Let $\boldsymbol{\beta}^*$ be the value of $\boldsymbol{\beta}$ so that the forecast error is *orthogonal* to or *uncorrelated* $\boldsymbol{X}_t$:

$$\mathrm{E}\left[\boldsymbol{X}_t \underbrace{\left(Y_{t+1} - \boldsymbol{X}_t'\boldsymbol{\beta}^*\right)}_{\text{forecast error}}\right] = \mathbf{0}.$$

- This is a *system* of equations.
- $\boldsymbol{\beta}^*$ minimizes the MSE.

Linear Projection

We can use the same steps as before to show that $\boldsymbol{\beta}^*$ minimizes MSE.

- Begin with an arbitrary linear forecasting rule,

$$Y^*_{t+1|t} = \boldsymbol{X}_t'\boldsymbol{\gamma}.$$

- Show that

$$\begin{aligned} MSE\left(Y^*_{t+1|t}\right) &= \mathrm{E}\left[\left(Y_{t+1} - \boldsymbol{X}_t'\boldsymbol{\gamma}\right)^2\right] \\ &= \mathrm{E}\left[\left(Y_{t+1} - \boldsymbol{X}_t'\boldsymbol{\beta}^* + \boldsymbol{X}_t'\boldsymbol{\beta}^* - \boldsymbol{X}_t'\boldsymbol{\gamma}\right)^2\right] \\ &= \mathrm{E}\left[\left(Y_{t+1} - \boldsymbol{X}_t'\boldsymbol{\beta}^*\right)^2\right] + \mathrm{E}\left[\left(\boldsymbol{X}_t'\boldsymbol{\beta}^* - \boldsymbol{X}_t'\boldsymbol{\gamma}\right)^2\right]. \end{aligned}$$

- Hence, MSE is minimized when $\boldsymbol{\gamma} = \boldsymbol{\beta}^*$.

Linear Projection

$Y^*_{t+1|t} = \boldsymbol{X}'_t\boldsymbol{\beta}^*$ is referred to as the *linear projection* of Y_{t+1} on $\boldsymbol{X}_t$.

- We will denote the linear projection as

$$\hat{P}(Y_{t+1}|\boldsymbol{X}_t) = \boldsymbol{X}'_t\boldsymbol{\beta}^* \quad \text{or} \quad \hat{Y}_{t+1|t} = \boldsymbol{X}'_t\boldsymbol{\beta}^*.$$

- Clearly

$$MSE\left(\hat{P}(Y_{t+1}|\boldsymbol{X}_t)\right) \geq MSE\left(\mathrm{E}\left[Y_{t+1}|\boldsymbol{X}_t\right]\right).$$

Linear Projection Solution

Using the orthogonality condition:

$$\boldsymbol{\beta}^* = \mathrm{E}\left[\boldsymbol{X}_t \boldsymbol{X}_t'\right]^{-1} \mathrm{E}\left[\boldsymbol{X}_t Y_{t+1}\right]. \tag{1}$$

- Least squares projection is the sample analogue of Equation (1).

Linear Projection MSE

Using our solutions for $\boldsymbol{\beta}^*$, we can solve for the MSE of the linear projection:

$$\begin{aligned}
MSE(Y^*_{t+1|t}) &= \mathrm{E}\left[\left(Y_{t+1} - \boldsymbol{X}'_t\boldsymbol{\beta}^*\right)^2\right] \\
&= \mathrm{E}\left[Y^2_{t+1}\right] - 2\mathrm{E}\left[Y_{t+1}\boldsymbol{X}'_t\boldsymbol{\beta}^*\right] + \mathrm{E}\left[\boldsymbol{\beta}^{*'}\boldsymbol{X}_t\boldsymbol{X}'_t\boldsymbol{\beta}^*\right] \\
&= \mathrm{E}\left[Y^2_{t+1}\right] - 2\mathrm{E}\left[Y_{t+1}\boldsymbol{X}'_t\right]\mathrm{E}\left[\boldsymbol{X}_t\boldsymbol{X}'_t\right]^{-1}\mathrm{E}\left[\boldsymbol{X}_tY_{t+1}\right] \\
&\qquad + \mathrm{E}\left[Y_{t+1}\boldsymbol{X}'_t\right]\mathrm{E}\left[\boldsymbol{X}_t\boldsymbol{X}'_t\right]^{-1}\mathrm{E}\left[\boldsymbol{X}_t\boldsymbol{X}'_t\right] \\
&\qquad\qquad \times \mathrm{E}\left[\boldsymbol{X}_t\boldsymbol{X}'_t\right]^{-1}\mathrm{E}\left[\boldsymbol{X}_tY_{t+1}\right] \\
&= \mathrm{E}\left[Y^2_{t+1}\right] - \mathrm{E}\left[Y_{t+1}\boldsymbol{X}'_t\right]\mathrm{E}\left[\boldsymbol{X}_t\boldsymbol{X}'_t\right]^{-1}\mathrm{E}\left[\boldsymbol{X}_tY_{t+1}\right].
\end{aligned}$$

Vector Linear Projection

Let $\boldsymbol{Y}_{t+1}$ be an $(n \times 1)$ vector and $\boldsymbol{X}_t$ an $(m \times 1)$ vector.

- The linear projection of $\boldsymbol{Y}_{t+1}$ on $\boldsymbol{X}_t$ is

$$\hat{P}(\boldsymbol{Y}'_{t+1}|\boldsymbol{X}_t) = \hat{\boldsymbol{Y}}'_{t+1|t} = \boldsymbol{X}'_t\boldsymbol{\beta}^*.$$

  where $\boldsymbol{\beta}^*$ is the $(m \times n)$ matrix such that

$$\mathrm{E}\left[\boldsymbol{X}_t\left(\boldsymbol{Y}'_{t+1} - \boldsymbol{X}'_t\boldsymbol{\beta}^*\right)\right] = \mathbf{0}.$$

- As in the univariate case

$$\boldsymbol{\beta}^* = \mathrm{E}\left[\boldsymbol{X}_t\boldsymbol{X}'_t\right]^{-1}\mathrm{E}\left[\boldsymbol{X}_t\boldsymbol{Y}'_{t+1}\right].$$