Some Basic Properties of Linear Projections

Lawrence J. Christiano

Background

- Linear projections play an important role in time series analysis.
- These notes provide an introduction to projections and some of their basic properties.
- In two companion sets of notes, we use the projection theory to derive the Kalman filter and the Kalman smoother.

An Example

- Let the log wage rate, w, and log price level, p, be given by

$$\begin{aligned} w &= z + u \\ p &= z + v, \end{aligned}$$

  where u and v are uncorrelated with each other and with z. All have zero mean.

- Suppose you observe w, but what you're really interested in is $w - p$.
  - obviously a move in w that reflects z is not interesting to you.

- You form the projection,

$$P\left[w - p | w\right] \equiv \alpha w,$$

  where α solves

$$\min_{\alpha} E\left[w - p - \alpha w\right]^2$$

Orthogonality Property of Projections:

- The optimization that defines the previous projection:

$$\min_{\alpha} E\left[w - p - \alpha w\right]^2$$

- First order condition after differentiating w.r.t. α :

$$E\overbrace{[w - p - \alpha w]}^{\text{projection error}} w \quad \overbrace{=}^{\text{orthogonality property}} \quad 0$$

solving this, one gets the familiar expression:

$$\begin{aligned}
\alpha &= \frac{E\left(w - p\right)w}{Ew^2} \\
&= \frac{E\left(u - v\right)\left(z + u\right)}{E\left(z + u\right)^2} \\
&= \frac{\sigma_u^2}{\sigma_u^2 + \sigma_z^2} = \frac{\sigma_u^2/\sigma_z^2}{\sigma_u^2/\sigma_z^2 + 1}
\end{aligned}$$

Orthogonality Property of Projections

- Suppose we have three random variables, Y, X_1 and X_2. We denote the projection of Y on a constant, X_1 and X_2, by

$$P[Y|1, X_1, X_2] = a_0 + a_1 X_1 + a_2 X_2,$$

where a_0, a_1, a_2 solve the projection problem:

$$\min_{a_0, a_1, a_2} E[Y - a_0 - a_1 X_1 - a_2 X_2]^2. \quad (1)$$

- If a_0, a_1, a_2 solve (1), then

$$\begin{aligned} E[Y - a_0 - a_1 X_1 - a_2 X_2] &= 0 \quad (2) \\ E[Y - a_0 - a_1 X_1 - a_2 X_2] X_1 &= 0 \\ E[Y - a_0 - a_1 X_1 - a_2 X_2] X_2 &= 0. \end{aligned}$$

- We now show that the reverse is also true:
  - if a_0, a_1, a_2 satisfy (2), then they solve (1).

Sufficiency of Orthogonality Condition

- Let

$$X = \left[\begin{array}{ccc} 1 & X_1 & X_2 \end{array} \right],$$

and let a denote the column vector of projection coefficients.

- We show: if the a's satisfy the orthogonality condition,

$$EX' [Y - Xa] = 0,$$

then those a's solve the projection problem so that

$$p [Y|X] = Xa.$$

- Note that satisfying the orthogonality condition is equivalent to

$$a = \left(EX'X\right)^{-1} EX'Y,$$

as long as there are no collinearities between the elements of X (i.e., $EX'X$ has full rank).

Sufficiency of Orthogonality Condition

- Consider an arbitrary column vector of numbers, g, and consider

$$E\left[Y - Xg\right]^2. \tag{3}$$

- Then,

$$\begin{aligned}
E\left[Y - Xg\right]^2 &= E\left[(Y - Xa) + X(a-g)\right]^2 \\
&= E\left[Y - a'X' + (a-g)'X'\right]\left[Y - Xa + X(a-g)\right] \\
&= E\left(Y - a'X'\right)(Y - Xa) + \overbrace{E(a-g)'X'X(a-g)}^{\geq 0} \\
&+ \overbrace{E\left[\left(Y - a'X'\right)X(a-g)\right] + (a-g)'EX'(Y - Xa)}^{=0 \text{ by orthogonality of } a}.
\end{aligned}$$

- The choice of g only affects the second term, which is minimized by $g = a$.
- Thus, if a satisfies the orthogonality condition, then $g = a$ optimizes (3).

Application #1 of Orthogonality Property

- Suppose we are given a particular linear combination of X

$$Xa,$$

having the property

$$EX'(Y - Xa) = 0$$

- Then, by sufficiency of orthogonality condition, we are entitled to assert:

$$P[Y|X] = Xa,$$

i.e., that a solves to the projection problem.

Application #2 of Orthogonality Property

- Suppose we have three zero-mean random variables, y, x, z, with

$$Eyx = Exz = 0, \text{ or, } x \perp y, \; x \perp z.$$

- Then,

$$P\left[y|x, z\right] = P\left[y|z\right].$$

- Proof: follows from sufficiency of orthogonality condition:

$$y - P\left(y|z\right) \perp x, z.$$

Because x, z are orthogonal to the projection error implied by $P\left(y|z\right)$, it follows that $P\left(y|z\right)$ is actually the projection of y on both x and z.

Application #3 of Orthogonality Property

- If $y \perp \xi$, then,

$$P[y|\xi] = 0.$$

- Proof follows from sufficiency of orthogonality condition:

$$E[y - 0]\xi = 0.$$

Recursive Property of Projections

- Consider the projection of Y onto a constant, X_1 and X_2 :

$$P[Y|1, X_1, X_2] = a_0 + a_1 X_1 + a_2 X_2,$$

so that

$$Y = a_0 + a_1 X_1 + a_2 X_2 + \varepsilon,$$

where $\varepsilon \perp X_1, X_2, 1$ (the last means, $E\varepsilon = 0$).

Now, project both sides onto $1, X_1$:

$$P[Y|1, X_1] = P[a_0 + a_1 X_1 + a_2 X_2 + \varepsilon|1, X_1]$$

$$\overbrace{=}^{\text{linearity of projections}} a_0 + a_1 X_1 + a_2 P[X_2|1, X_1] + \overbrace{P[\varepsilon|1, X_1]}^{=0 \text{ by application \#3}}$$

Recursive Property of Projections

- From previous expression,

$$\begin{aligned} Y - P\left[Y|1, X_1\right] &= a_0 + a_1 X_1 + a_2 X_2 + \varepsilon \\ &\quad - \left(a_0 + a_1 X_1 + a_2 P\left[X_2|1, X_1\right]\right) \\ &= a_2\left(X_2 - a_2 P\left[X_2|1, X_1\right]\right) + \varepsilon. \end{aligned}$$

- Since $\varepsilon \perp 1, X_1, X_2$, it follows that

$$a_2\left(X_2 - a_2 P\left[X_2|1, X_1\right]\right) = P\left(Y - P\left[Y|1, X_1\right]|1, X_1, X_2\right),$$

so

$$Y = P\left[Y|1, X_1\right] + P\left(Y - P\left[Y|1, X_1\right]|1, X_1, X_2\right) + \varepsilon.$$

Recursive Property of Projections

- The following two sets of information are the same:

$$\{1, X_1, X_2\} = \left\{ 1, X_1, \overbrace{X_2 - P\left(X_2|1, X_1\right)}^{\text{the part of } X_2 \text{ that cannot be determined from } 1, X_1} \right\}$$

so

$$\begin{aligned} & P\left(Y - P\left[Y|1, X_1\right]|1, X_1, X_2\right) \\ = \; & P\left(Y - P\left[Y|1, X_1\right]|1, X_1, X_2 - P\left(X_2|1, X_1\right)\right) \end{aligned}$$

- But,

$$\begin{aligned} 1, X_1 & \perp X_2 - P\left(X_2|1, X_1\right) \\ 1, X_1 & \perp Y - P\left[Y|1, X_1\right], \end{aligned}$$

so that by application #2

$$\begin{aligned} & P\left(Y - P\left[Y|1, X_1\right]|1, X_1, X_2\right) \\ = \; & P\left(Y - P\left[Y|1, X_1\right]|X_2 - P\left(X_2|1, X_1\right)\right). \end{aligned}$$

Recursive Property of Projections

- Then,

$$Y = P[Y|1, X_1] + P(Y - P[Y|1, X_1]|1, X_1, X_2) + \varepsilon.$$

implies:

$$Y = P[Y|1, X_1] + P(Y - P[Y|1, X_1]|X_2 - P(X_2|1, X_1)) + \varepsilon$$

- Since $\varepsilon \perp 1, X_1, X_2$, we get the *recursive property of projections*:

$$\begin{aligned} P[Y|1, X_1, X_2] &= P[Y|1, X_1] \\ &\quad + P(Y - P[Y|1, X_1]|X_2 - P(X_2|1, X_1)) \end{aligned}$$

- More generally:

$$P[Y|\Omega, X] = P[Y|\Omega] + P(Y - P[Y|\Omega]|X - P(X|\Omega))$$

Recursive Property of Projections

- The recursive property of projections says:

$$P[Y|\Omega, X] = P[Y|\Omega] + P(Y - P[Y|\Omega]|X - P(X|\Omega))$$

  - When you get new information, X, you adjust your previous guess, $P[Y|\Omega]$, by your best guess of the error in $P[Y|\Omega]$.
  - You make the adjustment using the part of X that really is new.

- Has interpretation as a learning algorithm, how a forecast is adjusted when new information comes in.

Law of Iterated Projections (LIP)

- According to LIP

$$P\left[P\left(Y|\Omega,X\right)|\Omega\right]=P\left[Y|\Omega\right].$$

- Proof: according to the recursive property of projections,

$$\begin{aligned} P\left[Y|\Omega,X\right] &= P\left[Y|\Omega\right]+P\left[Y-P\left(Y|\Omega\right)|X-P\left(X|\Omega\right)\right] \\ &= P\left[Y|\Omega\right]+\alpha'\xi \end{aligned}$$

where

$$\xi=\left[X-P\left(X|\Omega\right)\right]\perp\Omega$$

Now, project both sides on Ω :

$$P\left(P\left[Y|\Omega,X\right]|\Omega\right)=P\left[Y|\Omega\right]+\alpha'\overbrace{P\left[\xi|\Omega\right]}^{=0\text{ by application }\#3}$$