Quantitative Methods in Economics
Functional form

Maximilian Kasy

Harvard University, fall 2016

Roadmap, Part I

1. Linear predictors and least squares regression
2. Conditional expectations
3. **Some functional forms for linear regression**
4. Regression with controls and residual regression
5. Panel data and generalized least squares

Takeaways for these slides

Functional forms:

- Quadratic: decreasing or increasing returns
- Interactions: returns vary with covariates
- Discrete regressors, dummy variables, and saturated regressions
- Polynomial
- Linear in logarithms: elasticities
- Justification via Mincer model

- **Quadratic polynomial**:
- Y = earnings, Z = experience; $X_1 = Z$, $X_2 = Z^2$

$$E^*(Y \mid 1, X_1, X_2) = \beta_0 + \beta_1 X_1 + \beta_2 X_2$$

- Evaluating this at $Z = c$ gives $\beta_0 + \beta_1 c + \beta_2 c^2$.
- **Interactions**:
- Z_1 = experience, Z_2 = education; $X_1 = Z_1$, $X_2 = Z_2$, $X_3 = Z_1 \cdot Z_2$

$$E^*(Y \mid 1, X_1, X_2, X_3) = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3$$

- Evaluating this at $Z_1 = c$, $Z_2 = d$ gives $\beta_0 + \beta_1 c + \beta_2 d + \beta_3 c \cdot d$.

Questions for you

- Interpret these two functional forms.
- What happens as education is increased?
- How does that depend on the education we start with?
- How does that depend on experience?

- Recall conditional expectation: Solution to

$$\min_g E[Y - g(Z)]^2$$

is the *regression function*:

$$r(z) = E(Y \mid Z = z).$$

- Orthogonality Conditions: Consider any function $h(\cdot)$. Define

$$U = Y - r(Z).$$

- Then $U \perp h(Z)$, i.e. $E[Uh(Z)] = 0$, and in particular

$$E^*(Y|r(Z), h(Z)) = \beta_1 r(Z) + \beta_2 h(Z) = r(Z).$$

- Put differently: If $E(Y|X = x)$ is linear in x, then

$$E(Y|X = x) = E^*(Y|X = x).$$

Discrete regressors

- Assume
$$Z_1 \in \{\lambda_1, \dots, \lambda_J\}, \quad Z_2 \in \{\delta_1, \dots, \delta_K\}.$$
- Dummy Variables:
$$\begin{aligned} X_{jk} &= \begin{cases} 1, & if Z_1 = \lambda_j, Z_2 = \delta_k \\ 0, & otherwise \end{cases} \\ &= 1(Z_1 = \lambda_j, Z_2 = \delta_k). \end{aligned}$$
- Claim: $E(Y \mid Z_1, Z_2) = E^*(Y \mid X_{11}, \dots, X_{J1}, \dots, X_{1K}, \dots, X_{JK})$

Questions for you

Prove this.

Solution:

- Any function $g(Z_1, Z_2)$ can be written as

$$g(Z_1, Z_2) = \sum_{j=1}^{J} \sum_{k=1}^{K} \gamma_{jk} X_{jk}$$

  with $\gamma_{jk} = g(\lambda_j, \delta_k)$.
- Thus

$$E(Y \,|\, Z_1, Z_2) = \sum_{j=1}^{J} \sum_{k=1}^{K} \beta_{jk} X_{jk},$$

  where

$$\beta_{jk} = E(Y \,|\, Z_1 = \lambda_j, Z_2 = \delta_k).$$

- Since $E(Y|X = x)$ is linear in x, we get

$$E(Y \,|\, Z_1, Z_2) = E(Y|X) = E^*(Y|X).$$

Sample Analog

- Data: (y_i, z_{i1}, z_{i2}), $i = 1, \ldots, n$.
- Dummy Variables: $x_{i,jk} = 1(z_{i1} = \lambda_j, z_{i2} = \delta_k)$,

$$y = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}, \quad x_{jk} = \begin{pmatrix} x_{1,jk} \\ \vdots \\ x_{n,jk} \end{pmatrix}.$$

- Least Squares:

$$\min_b \sum_{i=1}^{n} \left(y_i - \sum_{j=1}^{J} \sum_{k=1}^{K} b_{jk} x_{i,jk} \right)^2$$

- This gives:

$$b_{jk} = \frac{\sum_{i=1}^{n} y_i x_{i,jk}}{\sum_{i=1}^{n} x_{i,jk}} = \bar{y} \,|\, \lambda_j, \delta_k.$$

Polynomial regressors

- Assume

$$E(Y|Z_1 = s, Z_2 = t) \cong \beta_0 + \beta_1 s + \beta_2 s^2 + \beta_3 t \cdot s + \beta_4 t + \beta_5 t^2.$$

- Example: Jacob Mincer, *Schooling, Experience and Earnings*, 1974, Table 5.1; 1 in 1000 sample, 1960 census; 1959 annual earnings; $n = 31093$;
- y = log(earnings), s = years of schooling, t = years of work experience;

$$\hat{y} = 4.87 + .255s - .0029s^2 - .0043t \cdot s + .148t - .0018t^2.$$

Predictive Effect

- Returns to college:

$$E(Y|Z_1 = 16, Z_2 = t) - E(Y|Z_1 = 12, Z_2 = t)$$
$$\cong \beta_1 \cdot 4 + \beta_2(16^2 - 12^2) + \beta_3 \cdot 4 \cdot t.$$

- Returns to high school:

$$E(Y|Z_1 = 12, Z_2 = t) - E(Y|Z_1 = 8, Z_2 = t)$$
$$\cong \beta_1 \cdot 4 + \beta_2(12^2 - 8^2) + \beta_3 \cdot 4 \cdot t.$$

Plugging in the estimates

Experience	Returns to college	Returns to high school
0	.70	.79
10	.52	.62
20	.35	.44

Questions for you

Verify this.

From predicting log W to predicting W

- Suppose $E(\log W \mid Z)$ is a linear function:

$$E(\log W \mid Z) = \beta_0 + \beta_1 Z.$$

- Define $U = \log W - \beta_0 - \beta_1 Z$, so $E(U \mid Z) = 0$.
-

$$W = \beta_0 + \beta_1 Z + U$$
$$\Rightarrow W = \exp(\beta_0 + \beta_1 Z) \cdot \exp(U).$$

- If U and Z are independent, $E[\exp(U) \mid Z]$ = $E[\exp(U)]$ and

$$\frac{E(W \mid Z = d)}{E(W \mid Z = c)} = \exp[\beta_1(d - c)] \cong \beta_1(d - c) + 1,$$

$$100\left[\frac{E(W \mid Z = d)}{E(W \mid Z = c)} - 1\right] \cong 100\beta_1(d - c).$$

Mincer model

- Compound Interest ($\triangle$ = fraction of one year):

$$\$1 \to \$(1 + r\triangle) \to \$(1 + r\triangle)^2 \to \$(1 + r\triangle)^3 \to \dots$$

- Annual Return $(1 + r\triangle)^{1/\triangle}$.

$$\lim_{x \to 0} \frac{f(1 + x) - f(1)}{x} = f'(1); \quad \lim_{x \to 0} \frac{\log(1 + x)}{x} = 1$$

$$\lim_{\triangle \to 0} \log[(1 + r\triangle)^{1/\triangle}] = r \cdot \lim_{\triangle \to 0} \frac{\log(1 + r\triangle)}{r\triangle} = r$$

- Thus

$$\lim_{\triangle \to 0} (1 + r\triangle)^{1/\triangle} = \exp(r) \approx 1 + r$$

- Works for small r:

$$\exp(.06) = 1.062; \quad \exp(.35) = 1.42; \quad \exp(.70) = 2.01.$$

- $PV(S)$ = present value at $t = 0$ of earning 0 while in school for an additional S years and then earning $W(S)$ for a very long time:

$$PV(S) = W(S) \int_S^\infty \exp(-rt)\,dt = W(S) \cdot \exp(-rS)/r.$$

- Returns such that students are indifferent about dropping out:

$$PV(S) = PV(0) \quad \Rightarrow \quad W(S) \cdot \exp(-rS)/r = W(0)/r,$$

- Thus:

$$\log(W(S)) = \log(W(0)) + rS.$$

- Linear Predictor:

$$E^*(Y \mid 1, S) = \gamma_0 + (r + \gamma_1)S,$$

where $Y = \log(W(S))$, $A = \log(W(0))$, and

$$E^*(A \mid 1, S) = \gamma_0 + \gamma_1 S.$$