

- Announcement: Papers due on Wednesday.
- Next week set aside for presentations. You decide amongst yourselves how to split up the presentation duties (everyone should say something though).
- We will do the presentations on Zoom. Too many students in quarantine.

## 1 Principal Components Analysis

1. PC analysis is one of several statistical factor analysis techniques.
2. Factor analysis: a method to explain the cross-sectional correlation of a large number of variables in terms of a smaller number of observed or unobserved variables. Fama-French 3 factor model, CAPM, these are observed factors. The econometrician has to take a stand on what the factors are. The statistical approach says the factors are unobserved and estimates them.
3. A two factor model for stock returns,  $r_{t,i}$ ,  $i = 1, \dots, n, t = 1, \dots, T$ , has the two-factor representation if

$$r_{t,i} = \alpha_i + \underbrace{\delta_{1,i}f_{1,t} + \delta_{2,i}f_{2,t}}_{\text{common part}} + \underbrace{\epsilon_{t,i}}_{\text{idiosyncratic}}$$

The slope coefficients  $\delta_{1,i}, \delta_{2,i}$  are called factor loadings. The factor loadings can (and will) differ across firms, according to their exposure to the (risk) factors.

The market compensates investors for bearing systematic risk, which is represented in the common part. Idiosyncratic risk can be diversified away by forming portfolios. No compensation for bearing idiosyncratic risk. And we start by assuming that  $f_{1,t}$  and  $f_{2,t}$  are unobserved.

4. The method of Principal Components: We make a couple of identifying assumptions to estimate the factors from returns  $r_{t,i}$ . Spoiler alert: the factors are going to be certain clever linear combinations of the returns data.
  - (a) There's no need to restrict ourselves to returns in computing the PCs. Example: each month, about 30 economic announcements. Can use the announcements as data and estimate factors (principal components) from these.

5. Write out the factor representation for each firm (ignore the constant)

$$\begin{aligned}
r_{t,1} &= \delta_{1,1}f_{t,1} + \delta_{2,1}f_{2,t} + \epsilon_{t,1} \\
r_{t,2} &= \delta_{1,2}f_{t,1} + \delta_{2,2}f_{2,t} + \epsilon_{t,2} \\
r_{t,3} &= \delta_{1,3}f_{t,1} + \delta_{2,3}f_{2,t} + \epsilon_{t,3} \\
&\vdots \\
r_{t,n} &= \delta_{1,n}f_{t,1} + \delta_{2,n}f_{2,t} + \epsilon_{t,n}
\end{aligned}$$

We estimate the factors (the principal components) sequentially. Estimate  $f_{t,1}$  first, then  $f_{t,2}$ . Let's assume a one-factor structure and write the equation system above in matrix form.

$$\begin{aligned}
r_{t,1} &= \delta_{1,1}f_{t,1} + \epsilon_{t,1} \\
r_{t,2} &= \delta_{1,2}f_{t,1} + \epsilon_{t,2} \\
r_{t,3} &= \delta_{1,3}f_{t,1} + \epsilon_{t,3} \\
&\vdots \\
r_{t,n} &= \delta_{1,n}f_{t,1} + \epsilon_{t,n}
\end{aligned}$$

$$\begin{aligned}
\underbrace{\begin{pmatrix} r_{1,1} & r_{1,2} & \cdots & r_{1,n} \\ r_{2,1} & r_{2,2} & & r_{2,n} \\ \vdots & & & \\ r_{T,1} & r_{T,2} & & r_{T,n} \end{pmatrix}}_r &= \underbrace{\begin{pmatrix} \delta_{1,1}f_{1,1} & \delta_{1,2}f_{1,1} & \cdots & \delta_{1,n}f_{1,1} \\ \delta_{1,1}f_{2,1} & \delta_{1,2}f_{2,1} & & \delta_{1,n}f_{2,1} \\ \vdots & & & \\ \delta_{1,1}f_{T,1} & \delta_{1,2}f_{T,1} & & \delta_{1,n}f_{T,1} \end{pmatrix}}_{\text{break apart}} \\
r &= \underbrace{\begin{pmatrix} f_{1,1} \\ f_{2,1} \\ \vdots \\ f_{T,1} \end{pmatrix}}_{f_1} \underbrace{\begin{pmatrix} \delta_{1,1} & \delta_{1,2} & \cdots & \delta_{1,n} \end{pmatrix}}_{\delta'_1} \\
r &= f_1 \delta'_1
\end{aligned}$$

(ignore the idiosyncratic part). The PC is not unique. Let  $c$  be some constant

$$r = f_1 \delta'_1 = (f_1 c) \left( \frac{\delta'_1}{c} \right)$$

So we normalize the factors such that  $\text{var}(f_1) = 1$ .

6. We want to find  $f_1$  and  $\delta_1$  that explains as much variation in  $r$  as possible. The sum of squares of  $(r - f_1 \delta'_1)$  (the unexplained part) is

$$\text{Tr}(r - f_1 \delta'_1)'(r - f_1 \delta'_1)$$

where  $\text{Tr}$  is the trace of the matrix, which is the sum of the diagonal elements. We choose  $f_1$  and  $\delta_1$  to minimize this thing. Let me show you. Let  $\tilde{r}_{t,i} = r_{t,i} - f_{t,1}\delta_{1,i}$ . and let  $T = 3$ ,  $n = 2$ .

(a)

$$\begin{aligned}
\text{Tr}(r - f_1\delta_1)'(r - f_1\delta_1) &= \text{Tr} \begin{pmatrix} \tilde{r}_{1,1} & \tilde{r}_{2,1} & \tilde{r}_{3,1} \\ \tilde{r}_{1,2} & \tilde{r}_{2,2} & \tilde{r}_{3,2} \end{pmatrix} \begin{pmatrix} \tilde{r}_{1,1} & \tilde{r}_{1,2} \\ \tilde{r}_{2,1} & \tilde{r}_{2,2} \\ \tilde{r}_{3,1} & \tilde{r}_{3,2} \end{pmatrix} \\
&= \text{Tr} \begin{pmatrix} \tilde{r}_{1,1}^2 + \tilde{r}_{2,1}^2 + \tilde{r}_{3,1}^2 & \tilde{r}_{1,1}\tilde{r}_{1,2} + \tilde{r}_{2,1}\tilde{r}_{2,2} + \tilde{r}_{3,1}\tilde{r}_{3,2} \\ \tilde{r}_{1,1}\tilde{r}_{1,2} + \tilde{r}_{2,1}\tilde{r}_{2,2} + \tilde{r}_{3,1}\tilde{r}_{3,2} & \tilde{r}_{1,2}^2 + \tilde{r}_{2,2}^2 + \tilde{r}_{3,2}^2 \end{pmatrix} \\
&= \tilde{r}_{1,1}^2 + \tilde{r}_{1,2}^2 + \tilde{r}_{2,1}^2 + \tilde{r}_{2,2}^2 + \tilde{r}_{3,1}^2 + \tilde{r}_{3,2}^2 \\
&= (\tilde{r}_{1,1} - f_{1,1}\delta_{1,1})^2 + (\tilde{r}_{1,2} - f_{1,1}\delta_{1,2})^2 + (\tilde{r}_{2,1} - f_{2,1}\delta_{1,1})^2 \\
&\quad + (\tilde{r}_{2,2} - f_{2,1}\delta_{1,2})^2 + (\tilde{r}_{3,1} - f_{3,1}\delta_{1,1})^2 + (\tilde{r}_{3,2} - f_{3,1}\delta_{1,2})^2
\end{aligned}$$

It's the sum of squared deviation of every observation from  $f\delta$ . Choose  $f$  and  $\delta$  to minimize this thing, and we get the first PC  $f_{t,1}$  and the factor loadings  $\delta_1$

7. To get the second PC and second set of factor loadings: Very simple. Repeat the above procedure, but using returns after controlling for the first factor. That is, replace  $r_{t,i}$  with  $r_{t,i} - f_{t,1}\delta_{1,i}$ , and define  $\tilde{r}_{t,i} = [(r_{t,i} - f_{t,1}\delta_{1,i}) - f_{t,2}\delta_{2,i}]$ , choose  $f_{t,2}$  and  $\delta_{2,i}$  to minimize the Trace of the analogous matrix. The result is

$$r_{t,i} = f_{t,1}\delta_{1,i} + f_{t,2}\delta_{2,i} + \epsilon_{t,i}$$

And because the factors are normalized,

$$\begin{aligned}
\text{var}(r_{t,i}) &= \delta_{1,i}^2 \text{var}(f_{t,1}) + \delta_{2,i}^2 \text{var}(f_{t,2}) + \text{var}(\epsilon_{t,i}) \\
&= \delta_{1,i}^2 + \delta_{2,i}^2 + \text{var}(\epsilon_{t,i})
\end{aligned}$$

The fraction of total variation in returns is the sum of squares of the factor loadings. Oh, I forgot to mention, the factors are chosen to be mutually uncorrelated.

8. In our dataset with  $n$  firms returns, we can compute  $n$  principal components. The procedure is useful to determine how many factors explain the data. After that, we might want to identify these unobserved factors with observable ones.