Macroeconometrics, Spring 2022
Bent E. Sørensen
February 20, 2022

1 Principal component.

Consider a collection of N or T dimensional column vectors X_k, where k indexes the variables, which can be investment, exports, ..., scores from a reading test, a math test, or whatever. Collect the vectors in a matrix X, and it sometimes very useful to construct the linear combination of the variables that captures most of the variation in X. (Because we focus on variation, assume all columns has mean zero.) We want weights, collected in a vector w^1 such that the elements sum to unity: $\Sigma_j w_j^1 = 1$, that maximizes

$$Var(Xw^1)$$

or

$$E\{w^{1\prime} X' X w^1\}.$$

We will not prove it, but it turns out that this problem has an elegant solution; namely, the optimal w^1 is the eigenvector of $X'X$ with the largest eigenvalue λ^1, and that eigenvalue is the maximum variance.

We call Xw^1 the first principal component and the interpretation is that it captures the most of the variation in the collection of variables. This may sound a little foggy, but that is because it is. But it has many applications. Think of a sample of people for which you have done a series of cognitive tests. How fast they read, how they can solve math puzzles, find a path out of a maze, etc. Now you want an estimated of peoples IQ. What is IQ? Not anything well-defined actually, but it is many times useful (for example, it is used a draft boards all over the world) to have some summary measure of cognitive ability. So often the largest principal component is used a measure of IQ.

But one principal component may not tell us all. So take the orthogonal part (to the first principal component) of all the variables, and now find the principal component again and call it Xw^2. It turns out that eigenvectors of a symmetric matrix are orthogonal, so actually we already have w^2 if we used an algorith to find the eigenvectors of $X'X$, and as the ws are orthogonal. so are the principal components. You can continue doing this and find p principal components and their eigenvalues. Normalize those to $\Sigma_1^p \lambda^i = 1$ and we interpret the λ^j's as measuring the fraction of variance "explained" by principal component j. In some applications, you may interpret the principal component as something economic, although as far as the math is concerned, it is just some linear combination. Example, look at world interest rates. Is most of the variation captured by one principal component? Maybe interpret this as the U.S. Fed (although you could test that). In the IQ case, people like to think that there is something that can be called general inteligence, but who knows? You may also use principal components to examine simply if a sample of variables are mainly capturing one underlying trend (in which case the larges eigenvalue captures most of the variance), that is what I did in my paper on county-level consumption in the 00's which is on my web-page.

The weights that each variable k gets in the linear combination that forms the first (second, third,...) principal component are called **loadings** and the loadings are often of interest. You may want to interpret your principal components. Say, you have a bunch of test scores, maybe it looks like that mathematical ones "load on" on one factor and the more verbal on on another. You might then conclude that intelligence come in the form of either mathematical and verbal (or, if three components are large: general, mathematical, and verbal). Many such examples exist. In economics, we might look for monetary (or demand) shocks and supply (or productivity) shocks. A typical application may first look for how many important principal components there is and then look at loadings for interpretation.

Principal components analysis is often used for descriptive purposes in economics and it can be hard to fit into a more rigorous testing framework. For example, the previous paragraph talked about "important PCs," but who decides what is important? It is still a useful tool, but it is often less suitable for testing which is why we don't priotitize teaching it in the econometrics sequence.