Coding Project 3: Principal Component Analysis of a Mass-on-a-spring System

Justin Hexem

Abstract

In this project, we will see if we can track a mass on a spring with varying levels of disturbances. This was done with three different camera angles for each of the four different experiments. For each of the experiments, we wish to analyse the vertical dynamics of this spring system. We would like to see if we can still detect the simple harmonic motion of the spring system through three different types of noise by using principal component analysis.

1 Introduction

Principal component analysis has a variety of different applications where it can help with simplifying large amounts of data. It can be used to help create a way of measuring general human intelligence by taking 56 different ways of measuring intelligence and eliminating many of the redundancies. The biggest principal component that arises from this analysis is much larger than the rest of the components, and we now call this component the Intelligence Quotient, or "IQ". Other applications of principal component analysis include quantitative finance, neuroscience, and population genetics to name a few.

Our application of principal component analysis for this project will be analyzing the dynamics of a mass-on-spring system. This system consists of an object with mass m that is attached to one end of a spring. The ODEs that govern these dynamics are given by

$$\frac{d^2 f(t)}{dt^2} = -\omega^2 f(t) \tag{1}$$

where $f(t)$ measures the displacement of the mass in the z-direction as a function of time. This ODE has the solution

$$f(t) = Acos(\omega(t) + \omega_0) \tag{2}$$

where A and ω_0 depend on the initial state of the system. We observe that this system can be completely characterized by the use of only one degree of freedom.

In this project, we will be analyzing data obtained from three different video cameras recording a mass-on-spring system from three different angles. This gives us many redundancies in our system that we would like to remove in order to better analyze the movement of the mass in the z-direction, and we can do this by making use of principal component analysis. We are taking our original data that has three degrees of freedom and condensing it into a system that only has one degree of freedom, which lines up with the ODEs that describe this system mentioned above. Principal component analysis can also be used to remove noise from our camera data by eliminating noise and any excess movement in the mass that is not in the z-direction.

2 Theoretical Background

We will now examine some mathematical concepts that are necessary for understanding how we can isolate the movement of the mass on the spring from the three different camera angles. The concepts that will help us understand this analysis include singular value decomposition, rank-n approximations, and principal component analysis.

2.1 Singular Value Decomposition

For any mxn matrix A, we can write A as the following product $A = U\Sigma V^*$ where U is a mxm unitary matrix, Σ is a mxn diagonal matrix, and V is a nxn unitary matrix. A matrix is a unitary matrix if the inverse of that matrix is equal to its conjugate transpose. The two unitary matrices can be geometrically interpreted as rotation matrices, and the diagonal matrix can be interpreted as a matrix that stretches vectors. We denote the diagonal entries the singular values of our original matrix A.

2.2 Rank-n Approximation

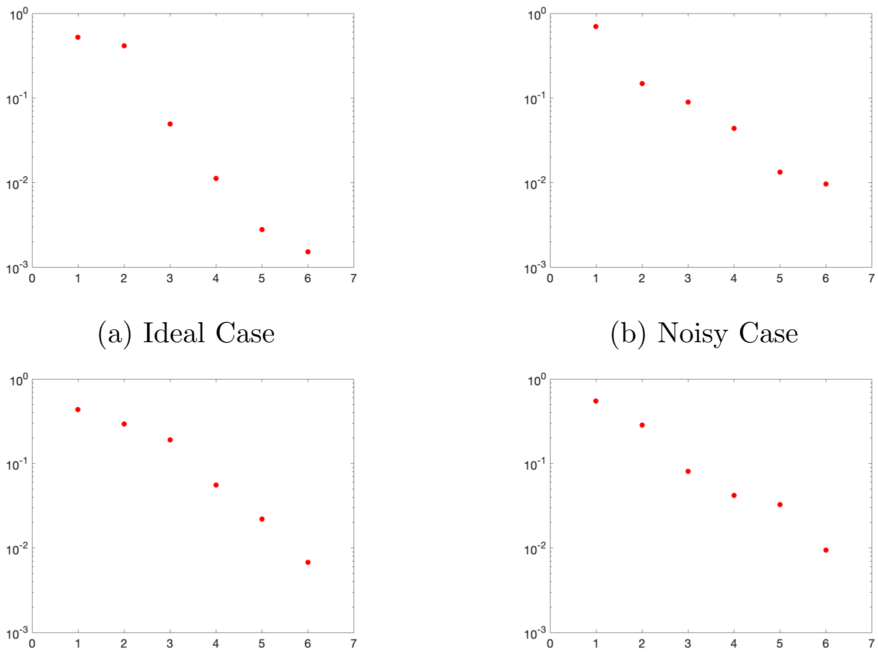
We may also write our mxn matrix as a sum of rank-one matrices in the following way: $A = \sum_{j=1}^{r} \sigma_j u_j v_j^*$. Writing our original matrix A in this way allows us to approximate A by only using some of the terms in this sum to obtain an approximation of A. Not only is this an approximation for A, but taking the first n terms of this sum gives the best possible n-term approximation of A. The nth partial sum of this series of rank-one matrices is called the rank-n approximation of A.

2.3 Principal Component Analysis

Principal component analysis is the process of finding a low-dimensional representation for data that is large, complex, and seemingly random. It involves standardizing our data, calculating the covariance matrix of our data, taking the singular value decomposition of this matrix, performing a coordinate transform to our original data matrix, and calculating rank-n approximations of our original data matrix. If we discover that a low-rank approximation of our data seems to represent our data extremely well, then we have removed the redundancy from our system and have a much simpler representation of our data. By examining the magnitude of our singular values, we can determine how much energy each of our rank-n approximations captures from our original matrix.

3 Numerical Methods

We will now discuss how we have handled the ideal case, since the other three cases can be handled with the same process. Principal component analysis will be used to achieve our results. We will first obtain the locations of the center of mass for each frame of each of the three camera angles. Once this is done, we will load the data for each camera angle into MATLAB as a matrix with two rows and a differing number of columns. In order to combine all of our data into one matrix, we must first make sure the number of columns in each of our three camera angles is the same. This can be achieved by truncating our data to match the number of columns in the matrix with the least amount of columns. Once this is done, we must subtract our data by the mean of each row of data and vertically stack our data into one matrix with six rows. Taking the singular value decomposition of this matrix will



(a) Ideal Case

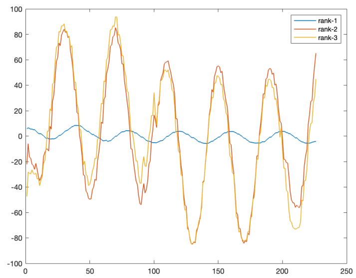
(b) Noisy Case

(c) Off Center Case

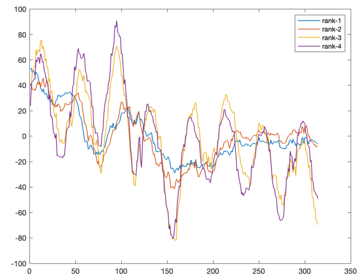
(d) Rotation Case

Figure 1: Singular Value Energies

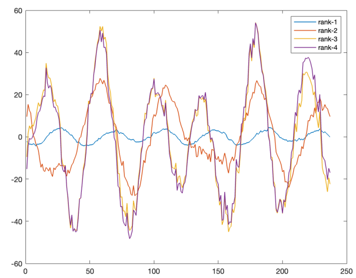
allow us to eliminate the redundant data from all of our measurements. We can then perform a coordinate transform on our original matrix using this singular value decomposition to obtain a new representation of our data along its principal components. We will then calculate the energies of each of our principal components by taking the square of each entry in our diagonal matrix and divide each of them by the sum of the squares of these entries. To come up with a lower dimensional approximation of our data, we will select the first n columns of our matrix U, the first n rows and columns of our matrix S, and the first n rows of our V^* matrix. This will give us a rank-n approximation of our data, and we can analyze which value of n gives us the best approximation of our data. This process is then repeated for the other three cases.



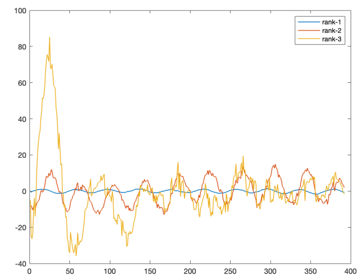
(a) Ideal Case



(b) Noisy Case



(c) Off Center Case



(d) Rotation Case

Figure 2: Low-Rank Approximation Convergence

4 Results

Figure 1 shows the energies of each of the six singular values for each of the four cases in a log plot. We see that in the ideal case that the singular values decline the fastest out of all of the plots, which makes sense because it has the least amount of noise. This implies the underlying system can be seen without needing too many terms for an accurate low-rank approximation. The off center and rotation cases seem to drop off fairly slowly, which can be seen by the later singular values having higher energies than in the other two cases. These lower values usually signify the noise in our image, so it would make sense in the latter two cases that this noise would be larger than in the ideal case.

Figure 2 shows the plots of the first few low-rank approximations for one of the camera angles in each case. The blue line is the rank-1 approximation, the red line is the rank-2 approximation, the yellow line is the rank-3 approximation, and the purple line is the rank-4 approximation. We can see that in the first three cases that as the approximation gets better, we converge to the underlying solution of the spring-mass system. The rotation case seems to converge to the same solution as the first three for the first two low-rank approximations. However, we can see that in the rank-3 approximation it looks like an entirely different periodic function. The first two singular values seem to pick up the overall movement of the mass on the spring, and the third singular value seems to pick up the rotation of the mass about the axis that runs through the spring. Overall, we can see that the oscillating frequencies for all four of the cases seem to be identical, which makes sense because they are all tracking the movement of the same system under different conditions. This shows us that our principal component analysis was able to successfully isolate the spring-mass system from our three different camera angles.

5 Conclusion

In this report, we were able to isolate the behavior of a spring-mass system from three different camera angles through different types of noise by making use of principal component analysis. We standardized our data so that we could create a single matrix that could hold all of our data, performed singular value decomposition to find the energies of our singular values, and used these energies to create low-rank approximations for our data that allowed us to extract the general behavior of the underlying system while removing most of the noise. This was possible because principal component analysis removed the redundancies and greatly reduced the dimensionality of our data so that we could isolate the movement of the mass oscillating on the spring.

I agreed with all of your choices of n for the rank-n approximations. It seemed like the values of n that I had chosen were trying to select the minimum rank approximation that picked up the general shape of a cosine solution with an amplitude that corresponded to our original data. This seemed to do the trick since including any higher-rank terms would only add in more noise, and this method of choosing n matched your choices exactly. It seemed like any noise that could arise from the way we tracked the mass was largely filtered out by our principal component analysis as well, so I do not think a more precise way of tracking the mass would make a significant difference here. I really enjoyed this project and it seemed very intuitive as long as I made sure I understood the underlying material well.