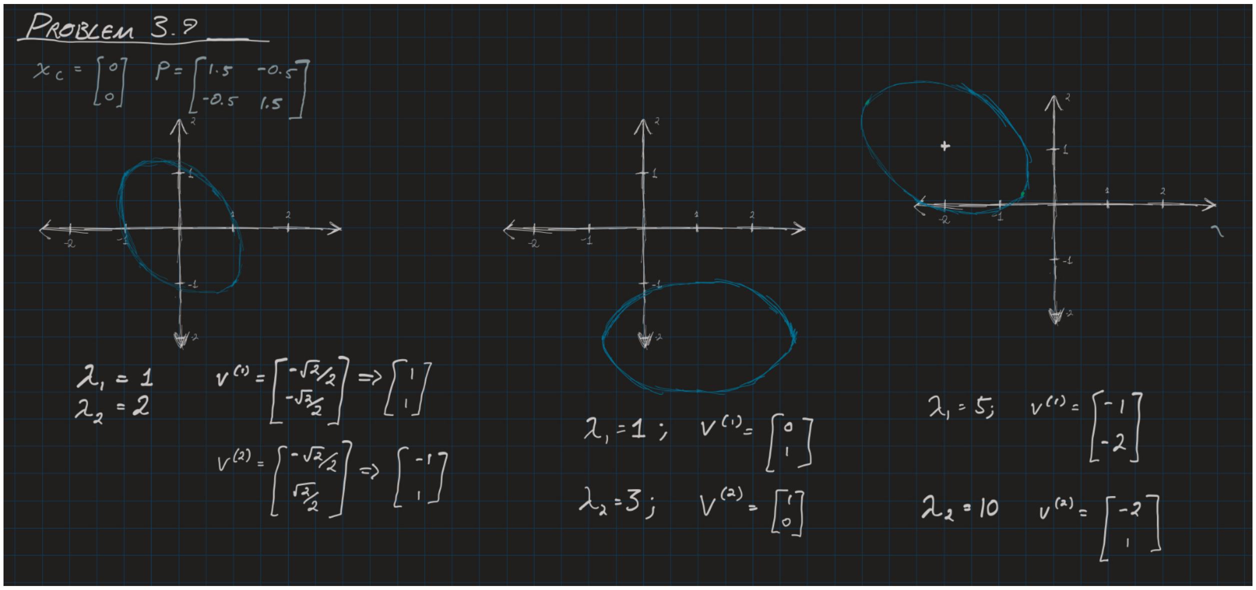
Problem Set 3 — ECE367

Aman Bhargava — 1005189733

3.9: Hand-Drawn Elipses



A useful interpretation of a zero-determinant matrix having one or more eigen values equal to zero is that, when transformed by A, a volume is 'squished' down to a lower dimension (and therefore, information is lost along one or more dimensions due to parts of the Λ term in $A = U\Lambda U^T$ being zero). The volume is therefore equal to zero as it lacks one or more dimensions within the relevant space.

3.10: Latent Semantic Indexing

1. `OptM problem 5.5 part 3`: $\tilde{M} = U\Sigma V^t$. What are interpretations of vectors u_l, v_l?

   One could interpret vectors u_l as how well word (represented by the indexes of u_l) relates with each of the r abstract 'concepts' extracted via the decomposition. Each of the vectors v_l could be interpreted as how each document relates to a given 'concept'.

2. `OptM problem 5.5 part 4`: We utilize the latent semantic space encoding of input vector x as $s_i = \Sigma^{-1} U^T x_i$. We compare similarity of vectors x_i, x_j based on the cosine of their angle $(s_i \cdot s_j)/(|s_i||s_j|)$.
3. Calculate $\tilde{M}$ and compute singular value decomposition. 10 Largest singular values:

```
Largest 10 singular values:
[1.5366294177331445, 1.0192424086695382, 0.958684541435874, 0.95391294599
51032, 0.9413064001927458, 0.9289078001291811, 0.8977405000640665, 0.8918
819220380092, 0.8686645393885041, 0.8160833878423517]
```

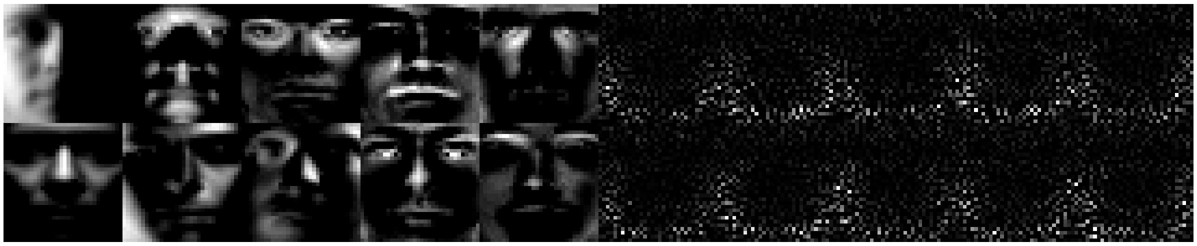
4. Two must similar documents: #10 and #9 (Barack Obama and George W. Bush).
5. Lowest k that preserves maximum similarity: $k = 3$. Most similar documents for $k = 2$ are #6 and #1 (B. J. Cole and John Holland (composer)).

3.11: Eigen Faces

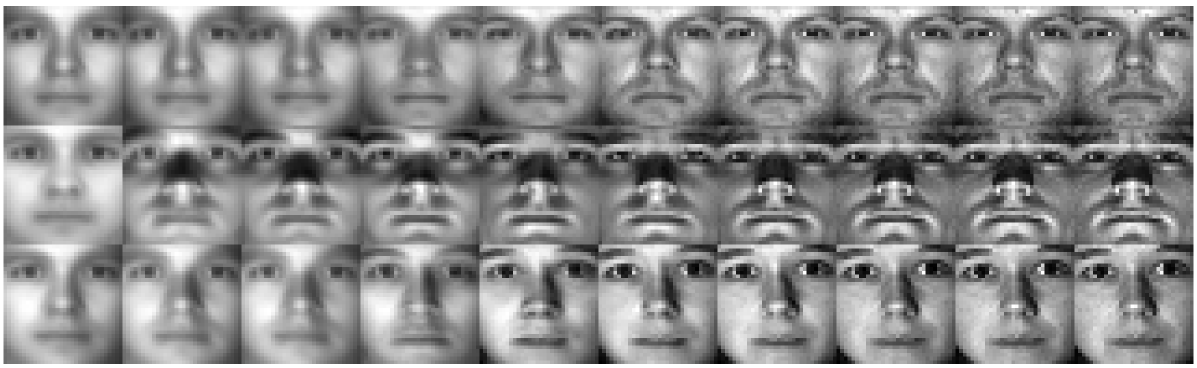
1. **Connection between singular values of X and eigendecomposition of C:**

   Since $C = XX^T$, the eigenvalues of C comprise the **left singular values** of X.
2. Since C is a positive definite matrix, its eigenvalues are real.
   <<Insert image of λ graph here>>
3. The 10 eigenfaces on the left are correspond to the 10 highest eigenvalues while the 10 eigenfaces on the right correspond to the 10 lowest eigenvalues.

   The left eigenfaces are more recognizable as components of the face as they account for a larger portion of the variance in each image. The right eigenfaces are less recognizable as they account for a small portion of the variance of what makes a picture of a face.



4. Each row of the following figure is one face (face #1, #1076, and #2043) at the requested 10 levels of detail (lowest to highest from left to right).



5. Below is the tabulated Euclidian distances between each face (indexes i, j refer within the concatenated set $\mathcal{I} = \{1, 2, 7, 2043, 2044, 2045\}$.

   Clearly the distance between vectors belonging to the same person have a lower mean.

   One could make a face recognition system by setting up a cutoff for the similarity euclidian distance between the projected image vectors for the first batch of principal components (e.g. 25).

```
IN GROUP:
i: 1; j: 2 => 2.324718264470165
i: 1; j: 3 => 1.8525992183972875
i: 2; j: 3 => 1.5381937356655304
i: 4; j: 5 => 2.9756750686724334
i: 4; j: 6 => 2.6714709875212983
i: 5; j: 6 => 3.24857261017142
-------------------------------
Mean:         2.43

OUT GROUP:
i: 1; j: 4 => 4.245540532925109
i: 1; j: 5 => 5.991480961254487
i: 1; j: 6 => 5.356871184687975
```

```
i: 2; j: 4 => 5.504999434433608
i: 2; j: 5 => 7.461832224375202
i: 2; j: 6 => 6.3229810893906215
i: 3; j: 4 => 4.956899681915511
i: 3; j: 5 => 7.022535289269725
i: 3; j: 6 => 6.19188745064867
-------------------------------
Mean:         5.89
```