Solutions to Homework Set Six
ECE 175
Electrical and Computer Engineering
University of California San Diego

Nuno Vasconcelos

1. a) Using the eigenvalue decomposition of $\mathbf{A}$

$$\mathbf{A} = \mathbf{\Phi}\mathbf{\Lambda}\mathbf{\Phi}^T$$

we can write

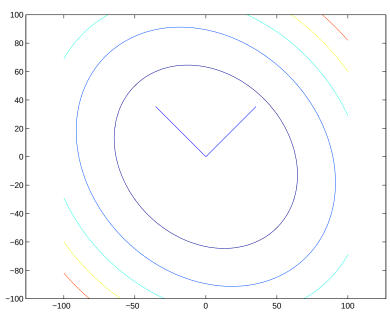
$$f(\mathbf{x}) = (\mathbf{\Phi}^T\mathbf{x})^T\mathbf{\Lambda}\mathbf{\Phi}^T\mathbf{x}$$

which, making $\mathbf{y} = \mathbf{\Phi}^T\mathbf{x}$, is identical to

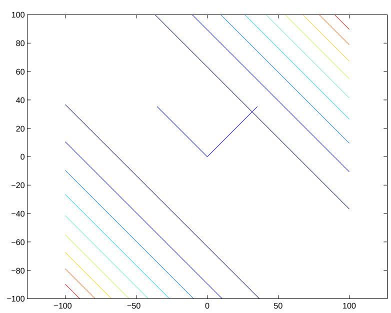
$$f(\mathbf{y}) = \mathbf{y}^T\mathbf{\Lambda}\mathbf{y} = \sum_i \lambda_i y_i^2.$$

Hence the transformation is $\mathbf{\Phi}^T$, the rotation by the matrix whose rows are the eigenvectors of $\mathbf{A}$, and the coefficients are the eigenvalues λ_i of $\mathbf{A}$.

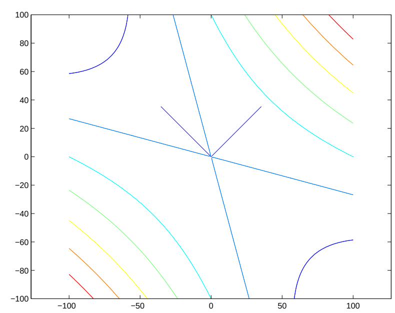
b) For this matrix, modifying the value of b does not change the eigenvectors, only the associated eigenvalues. Note that the eigenvectors define the principal axis of the surface $f(\mathbf{x})$.



$b = 1$



$b = 5$

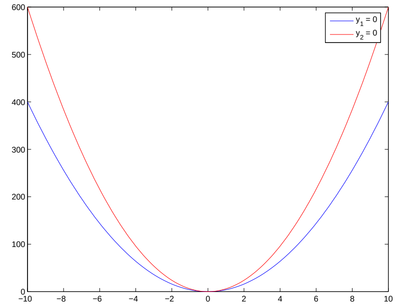


$b = 10$

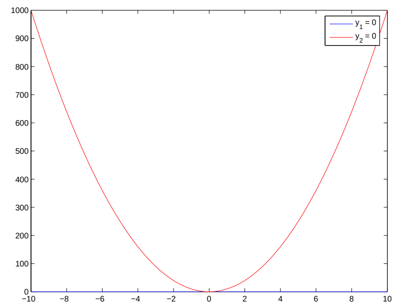
c) The eigenvalues are:

$b = 1$	$b = 5$	$b = 10$
$(6, 4)$	$(10, 0)$	$(15, -5)$

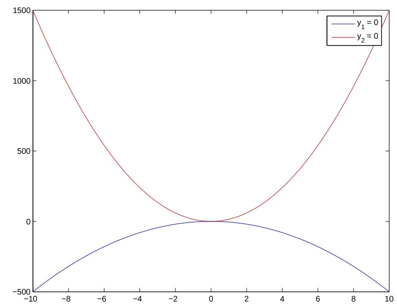
The plots of the corresponding 1D slices are shown below.



$b = 1$



$b = 5$



$b = 10$

d) The eigenvalues determine the curvature of $f(\mathbf{x})$ along the direction of the eigenvectors. When both eigenvalues are positive ($b = 1$), the function has positive curvature along the two eigenvector directions. Hence it is a bowl. A zero eigenvector ($b = 5$) means no curvature along the associated eigenvector. Hence the function is flat, or 1D. A negative eigenvalue ($b = 10$) means negative curvature along one of the eigenvectors. Hence the function is "cup-like" in one dimension, and "cap-like" in the second, i.e. a "saddle".

2.a) Last week we saw that the mixture has covariance

$$\begin{aligned}
\mathbf{\Sigma}_x &= E_{\mathbf{X}}[(\mathbf{x} - \mu_x)(\mathbf{x} - \mu_x)^T] \\
&= \sum_{i=1}^{C} \pi_i [\mathbf{\Lambda}_i + (\mu_i - \mu_x)(\mu_i - \mu_x)^T] \\
&= \mathbf{S}_x + \sum_{i=1}^{C} \pi_i \sigma_i^2 \mathbf{I}
\end{aligned}$$

Letting $\mathbf{\Phi}$ and $\mathbf{\Gamma}$ be the eigenvalue decomposition of $\mathbf{S}_x$, i.e.

$$\mathbf{S}_x = \mathbf{\Phi}\mathbf{\Gamma}\mathbf{\Phi}^T,$$

we can write

$$\begin{aligned}
\mathbf{\Sigma}_x &= \mathbf{S}_x + \sum_{i=1}^{C} \pi_i \sigma_i^2 \mathbf{I} \\
&= \mathbf{\Phi}\mathbf{\Gamma}\mathbf{\Phi}^T + \sum_{i=1}^{C} \pi_i \sigma_i^2 \mathbf{I} \\
&= \mathbf{\Phi}\mathbf{\Gamma}\mathbf{\Phi}^T + \sum_{i=1}^{C} \pi_i \sigma_i^2 \mathbf{\Phi}\mathbf{I}\mathbf{\Phi}^T \\
&= \mathbf{\Phi} \left(\mathbf{\Gamma} + \sum_{i=1}^{C} \pi_i \sigma_i^2 \mathbf{I} \right) \mathbf{\Phi}^T.
\end{aligned}$$

Note that the matrix in the brackets is diagonal, from which the principal components of $\mathbf{\Sigma}_x$ have identical direction to those of $\mathbf{S}_x$. The length of the i^{th} component of $\mathbf{\Lambda}_x$ is

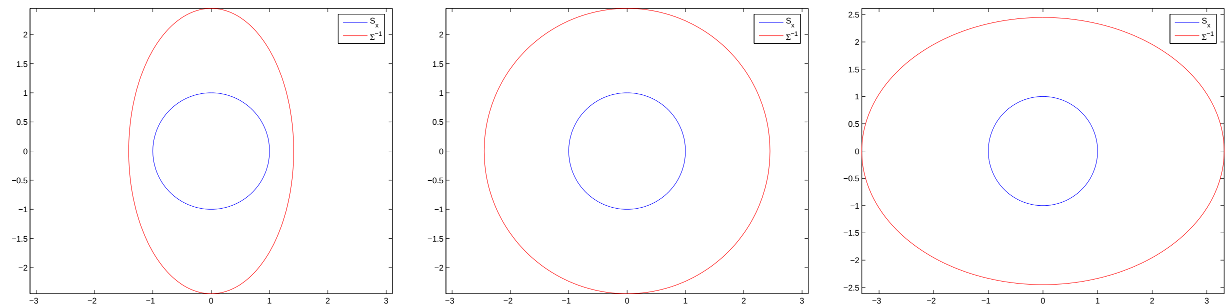
$$\lambda_i = \gamma_i + \sum_{k=1}^{C} \pi_k \sigma_k^2,$$

where γ_i is the length of the corresponding component of $\mathbf{\Gamma}$.

b) In this case,

$$\mathbf{\Sigma}_x = \mathbf{S}_x + \mathbf{\Lambda} = \begin{bmatrix} c+1 & 0 \\ 0 & 6 \end{bmatrix}$$

and the contour plots are as shown below.



$c = 1$ $c = 5$ $c = 10$

The components of $\mathbf{\Lambda}_i$ stretch the contours of the Mahalanobis distance of $\mathbf{S}_x$ in the horizontal and vertical direction.

3.a)

1.

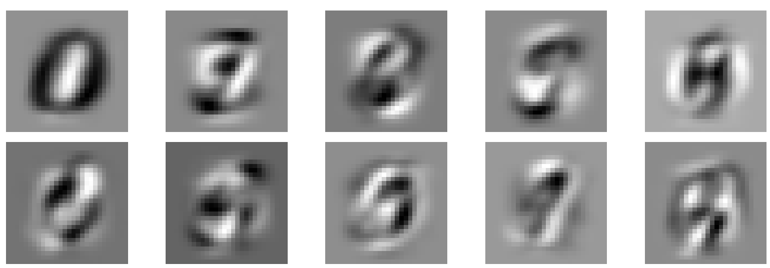


Figure 1: Top 10 eigendigits for the entire dataset

2. Classification results

   (a) The eigenvalues decrease rapidly, and beyond the 100^{th} eigenvalue the amplitudes are negligible as compared to the largest. Hence, a good guess would be to use less than 100 dimensions for the PCA subspace.

   (b) Classification Results - using very low-dimensional subspaces the error rates are quite high. The error is minimum around 30 dimensions and increases beyond that.

   (c) In problem set 3, without class covariance information and using all dimensions, we obtained an error rate of 0.22. This is much worse than using PCA with a 30 dimensional subspace, for which the error rate is 0.048. We can obtain the BDR by using `mvnpdf` function in Matlab, or by evaluating the multivariate Gaussian expression itself. Probably due to some

Figure 2: Top 10 eigendigits for class of digit 5

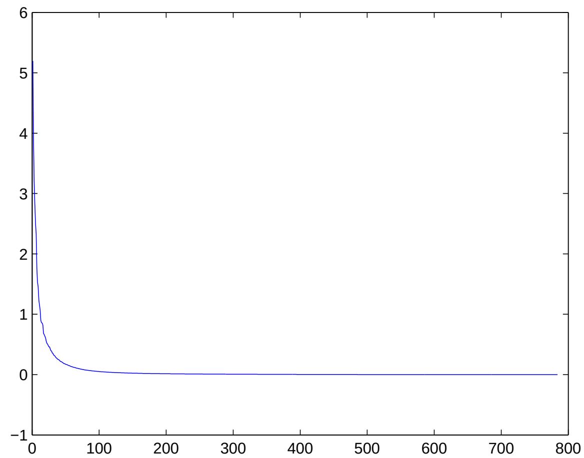


Figure 3: Eigenvalue magnitude across PCA component

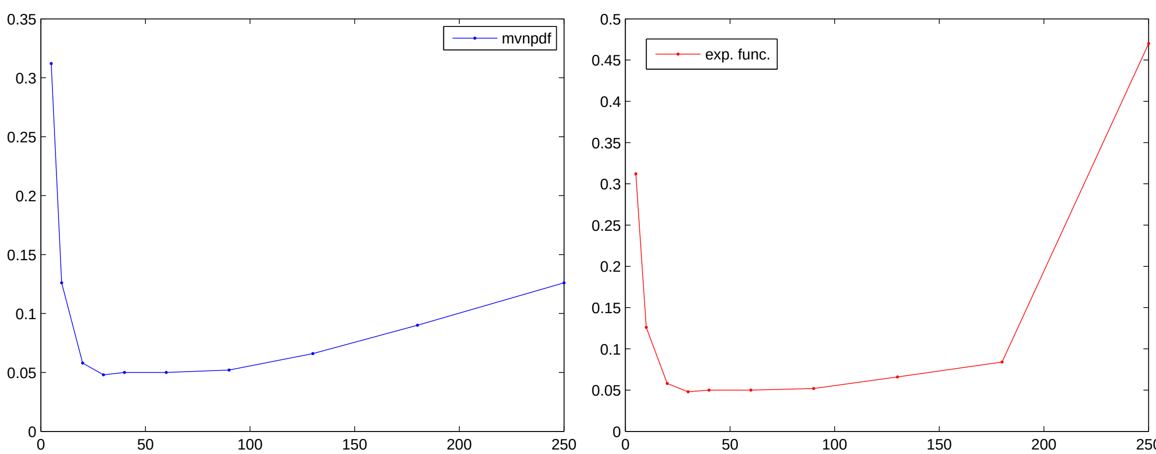


Figure 4: Error rates vs dimension: using `mvnpdf` (left) evaluating the multivariate Gaussian expression (right)

numerical instability in the computation of the inverse of Σ, the results degenerate after

180^{th} dimension.

3. Top 10 least-5 like images as obtained using the non-principal components of class of digit5

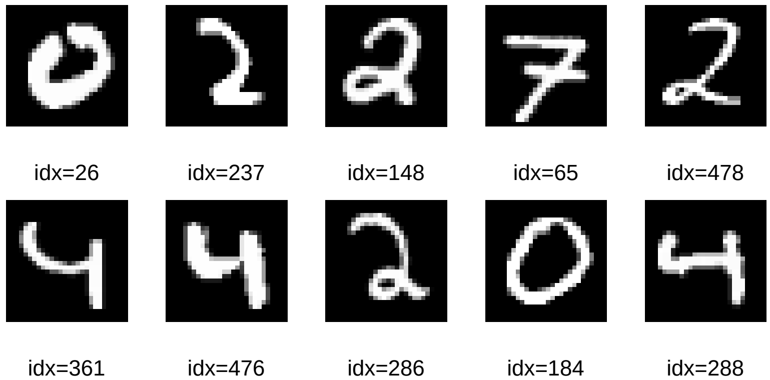


Figure 5: Least 5 like images