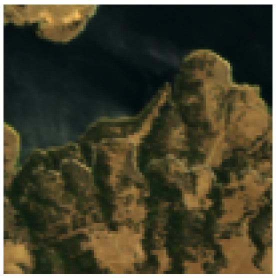
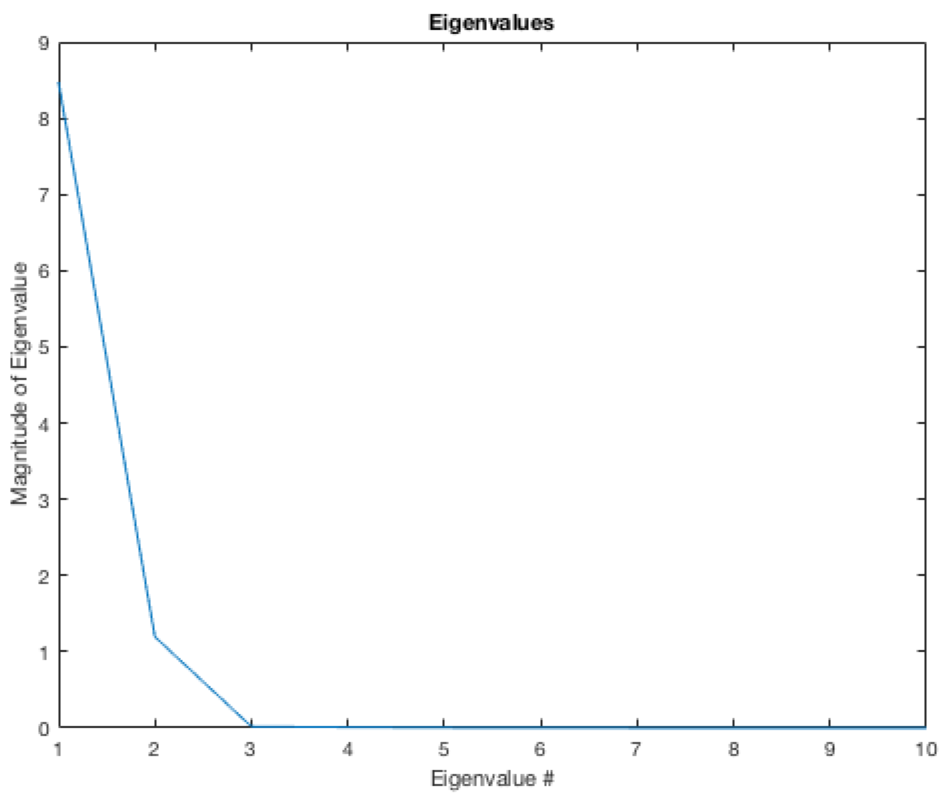
Table of Contents


```
load data_lab3
figure
format long
imshow(mat2gray(im(:,:,[30,20,7])));
```



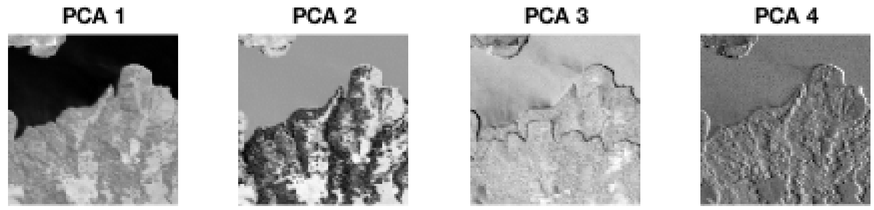
```
%{
1.
Perform a PCA on the data, seeing each pixel as an observation, and
 each
spectral band as a variable. Plot the eigenvalues in descending order,
 and
show a few of the corresponding PCA scores (the coordinates in the new
basis). Since you have coordinates for each pixel, you can display
 them as
images. Comment on the obtainedimages (salient features in the
 principal
components). Compare with the color composition you can make using the
 RGB
bands. Display the PCA scores for the last few components. How do you
interpret them?
%}

%reshaping our image into our matrix X' where the columns are spectra
 and the
%rows are pixels
img = zeros(10201, 195);
```

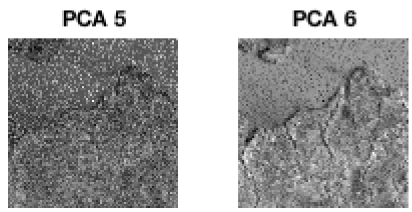
```
for i = 1:195
    img(:, i) = reshape(im(:, :, i) , 10201, 1);
end
co_variance = cov(img);
sample_mean = mean(img);
[V, D] = eigs(co_variance, 195);

%Sort eigenvalues in order
D_s = sort(diag(D),'descend');

%Plot of eigenvalues
%figure
x_axis = 1:10;
figure
plot(x_axis,D_s(1:10));
xlabel('Eigenvalue #')
ylabel('Magnitude of Eigenvalue')
title('Eigenvalues')
```

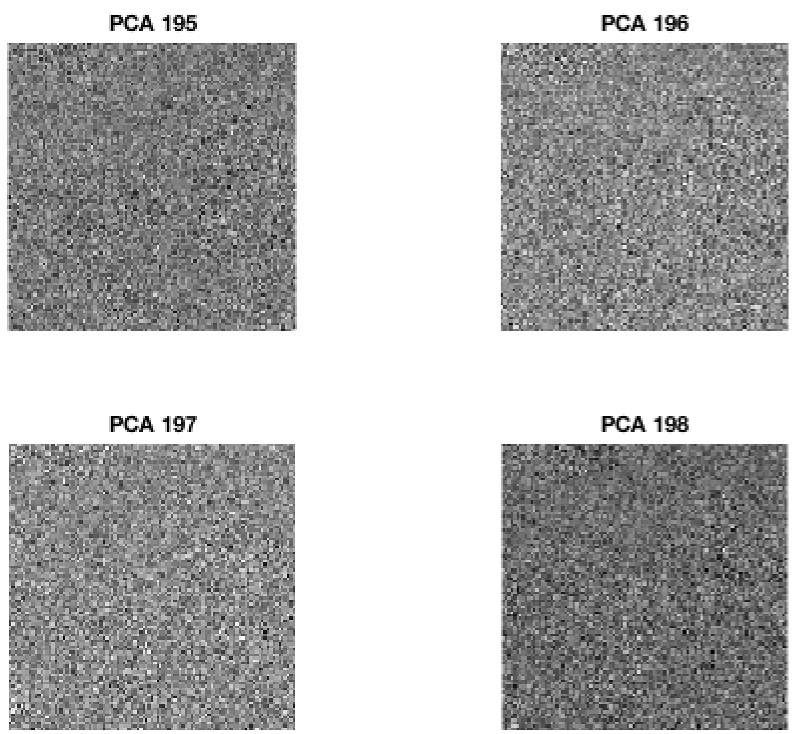



```
score = zeros(10201, 10);

%first five PCA scores
for i = 0:5
    score(:, i+1) = img*V(:, 195 - i);
end
```

```
figure
for i = 1:6
    subplot(2,4,i);
    imshow(mat2gray(reshape(score(:, i), 101, 101)));
    title(['PCA ' int2str(i)]);
end

%Last 4 PCA Scores
for i = 1:4
    score(:, i+6) = img*V(:, i);
end

figure
for i = 1:4
    subplot(2,2,i);
    imshow(mat2gray(reshape(score(:, i+6), 101, 101)));
    title(['PCA ' int2str(194+i)]);
end


%{
Comment: We see that the images of the PCA scores increasingly get
 worse
because the principal components associated with the scores capture
 less
and less of the variance in the data. The first couple of images of
 the
PCA scores look most similar to our image with the RGB bands, since
the principal components of those images captures most of the variance
in the data. Specifically, we can see the water and terrain are very
similar in our image with the RGB bands compared to the images of the
major PCA scores. Also, the images of the last 4 PCA scores are
 impossible
to decipher because they capture very little information in our data.

%}
```


PCA 1
PCA 2
PCA 3
PCA 4


PCA 5
PCA 6

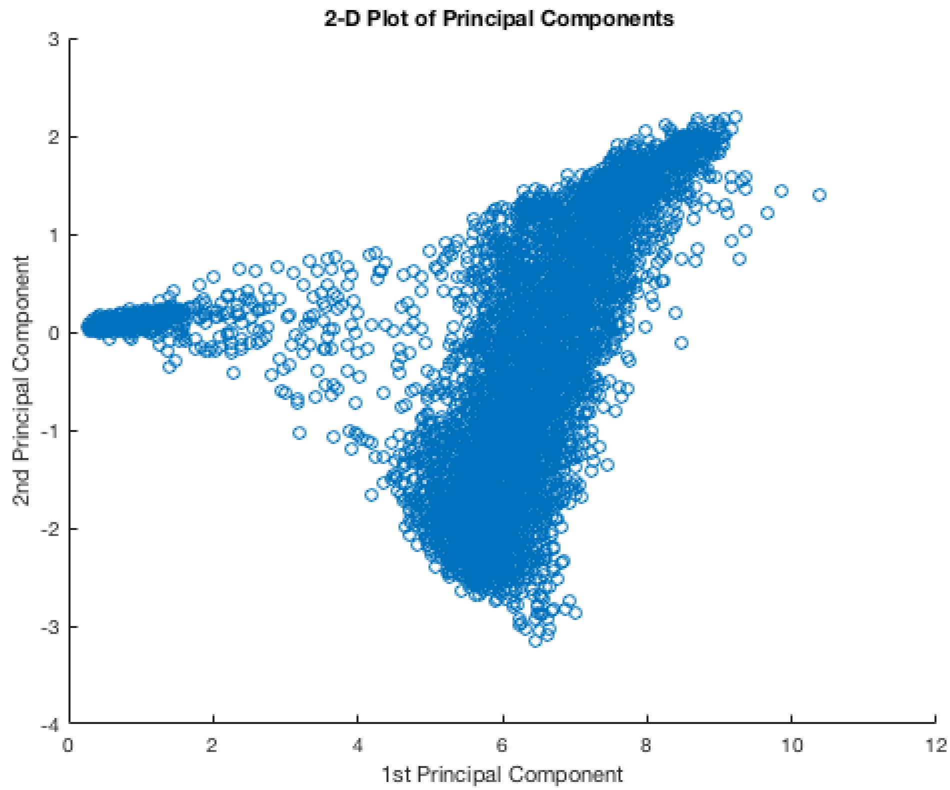

PCA 195
PCA 196
PCA 197
PCA 198

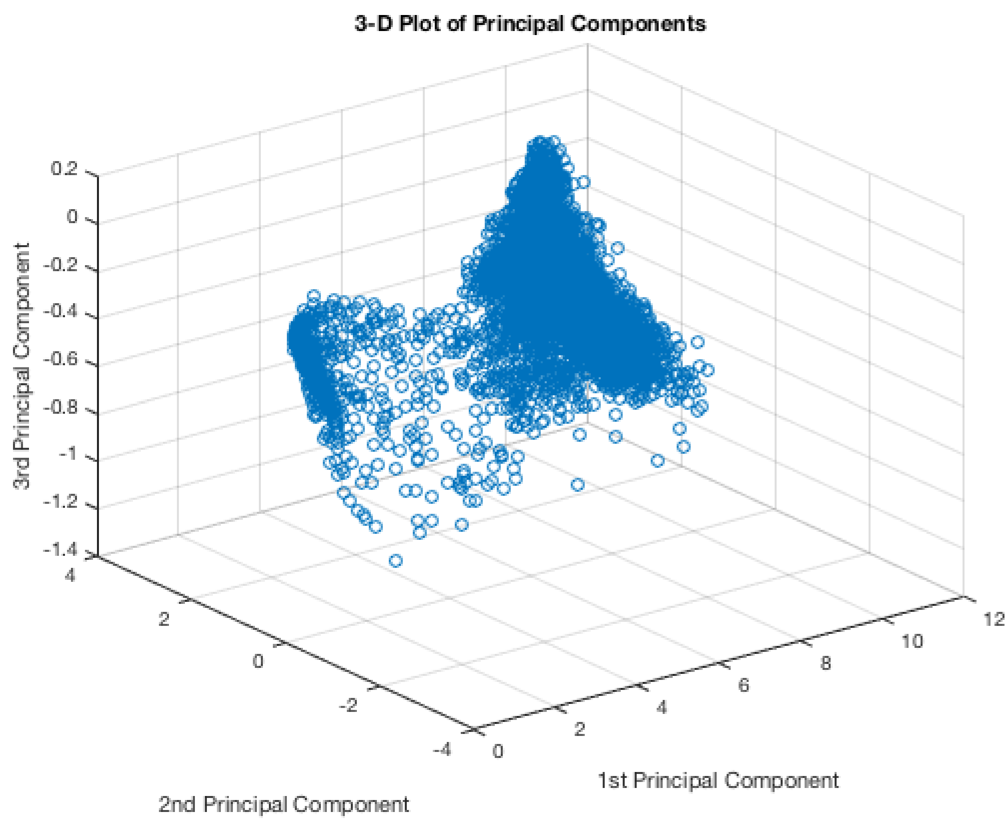
2.

```
%{
Based on our plot of the eigenvalues and our images of the PCA scores,
 we
can say that the subspace occupied our data has a dimension of 3. This
 is
because only the first three principal components are significant,
 which is
confirmed by our plot of the eigenvalues which go to 0 after the 3rd
eigenvalue.
%}
```

3.

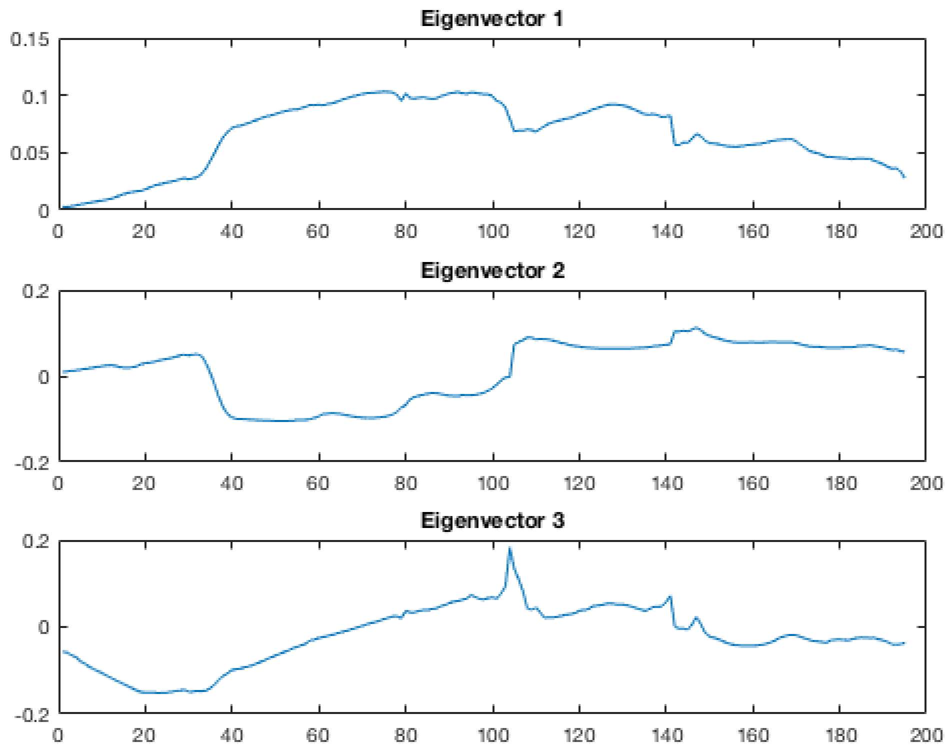
```
%2D-Plot of Principal Components
figure
scatter(score(:, 1), score(:,2));
%scatter3(score(:,1), score(:,2), score(:,3));
xlabel('1st Principal Component')
ylabel('2nd Principal Component')
title('2-D Plot of Principal Components ')


%3D-Plot of Principal Components
figure
scatter3(score(:,1), score(:,2), score(:,3));
xlabel('1st Principal Component')
ylabel('2nd Principal Component')
zlabel('3rd Principal Component')
title('3-D Plot of Principal Components')

%{
The plots do not confirm our answer to the previous problem because we
 see
2 clusters of points in our 2-D and 3-D plot. These clusters suggest
 that
we have 2 main regions where our points reside when we project them
 onto
our eigenbasis. This suggests that that our data has two main features
 or
2 dimensions.

%}
```

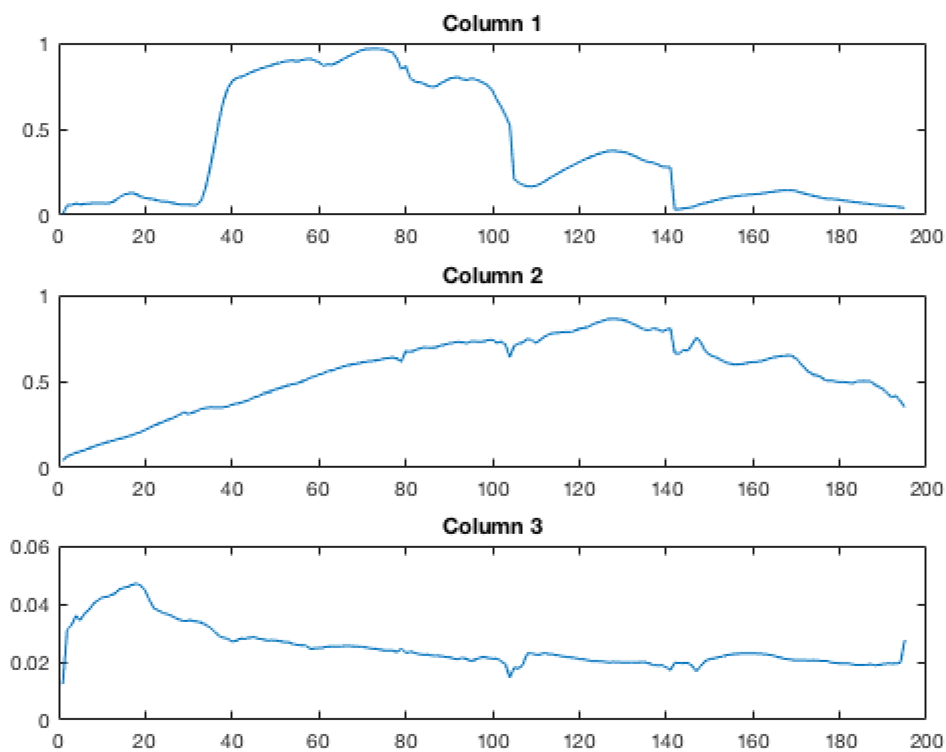

2-D Plot of Principal Components
1st Principal Component
2nd Principal Component
0
2
4
6
8
10
12
3
2
1
0
-1
-2
-3
-4


3-D Plot of Principal Components
3rd Principal Component
2nd Principal Component
1st Principal Component
0.2
0
-0.2
-0.4
-0.6
-0.8
-1
-1.2
-1.4
4
2
0
-2
-4
0
2
4
6
8
10
12

4.

```
figure
for i = 1:3
    subplot(3,1,i);
    plot(V(:, 195 - i + 1))
    title(['Eigenvector ' int2str(i)]);

end

figure
for i = 1:3
   subplot(3,1,i)
   plot(S(:,i));
   title(['Column ' int2str(i)]);
end

%{
While it has hard to notice any direct similarities, we see that the
 shapes
of the 1st and 2nd eigenvector have a similar shape to the 1st and 2nd
column of S. This would suggest that our eigenbasis closely represents
the spectra bands of two materials since those are the columns of S.
The reason why the third eigenvector is not as similar to the third
 column
of S may be due to the fact that the spectral bands of the 3rd
 material
are less abundant, so they are harder to represent.
%}

%5.
%{
If the linear model holds, then if we follow the solution to the
 model,
we see that the column space of S can represent the data. Since we
 have
3 independent columns of S (this is because S has to be full rank in
 order
to take the inverse of (S^t*S), the subpsace of our data is in a
3-dimensional space. Conceptually that makes sense because we have 3
 main
materials: soil, vegetation, and water. These three materials are
represented as spectral bands through the columns of S and we can
 represent
our data as a linear combination of the spectral bands of these
 materials
using the equation X = SA + N.

%}
```


Eigenvector 1
0.15
0.1
0.05
0
0 20 40 60 80 100 120 140 160 180 200
Eigenvector 2
0.2
0
-0.2
0 20 40 60 80 100 120 140 160 180 200
Eigenvector 3
0.2
0
-0.2
0 20 40 60 80 100 120 140 160 180 200


Column 1
1
0.5
0
0 20 40 60 80 100 120 140 160 180 200
Column 2
1
0.5
0
0 20 40 60 80 100 120 140 160 180 200
Column 3
0.06
0.04
0.02
0
0 20 40 60 80 100 120 140 160 180 200

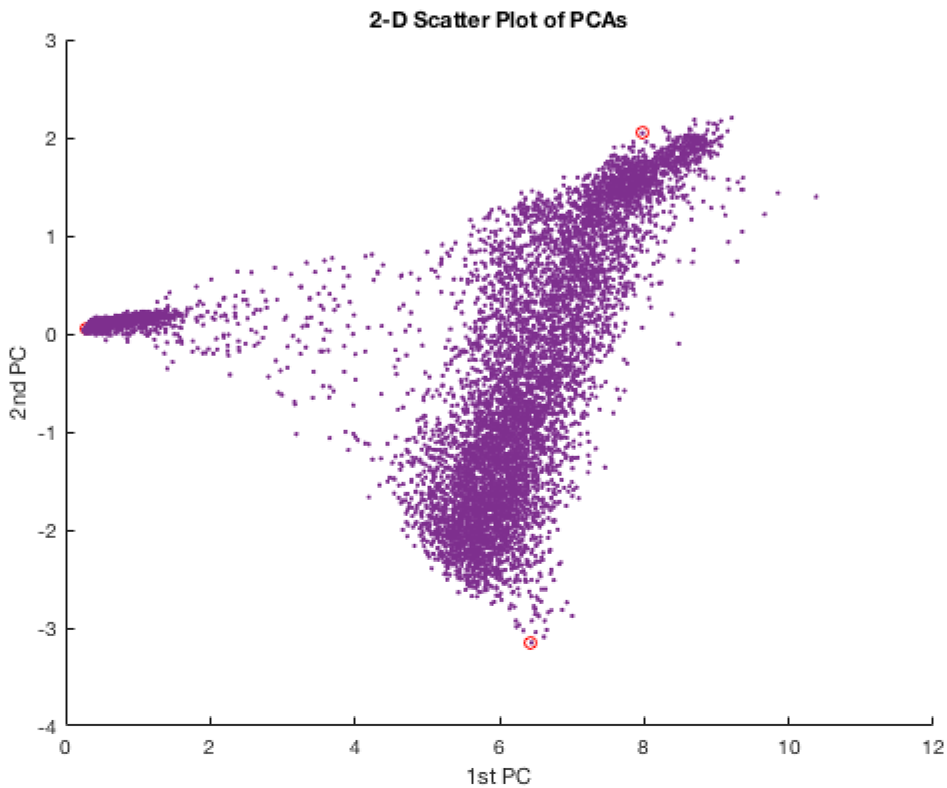
```
%{
6.
By introducing the sum to one constraint, we reduce the 3-D space
spanned the columns of S to dimensions because if we know the
 proportions
of two of our constraints(materials), we can calculate the third
 constraint
(material) as well. In other words, we can calculate the proportions
 of
the third material in each pixel by knowing the proportions of the
 other
two materials. This means that we only have 2 degrees of freedom.
 Thus,
our subspace is now reduced to 2 dimensions.
%}


%{
7.
The positivy constraint reduces the space of our 2-dimensional
 subspace,
but does not reduce the dimensionality of our subspace. We can think
 of
this as restricting the subpspace to only points that follow the
 positivy
constraint. We can think of our subpsace as a triangle where each
 corner
represents a material and all points inside the triangle represent
 points
that follow the sum-to-one and positivy constraint. This is confirmed
further by our 2-D scatterplot of our PCA scores, which has a
 triangular
shape.

%}

%{
8. The instrinsic dimensionality of our subspace is 2 because the
constraints placed on our linear model, reduce the degrees of freedom
 from
3 to 2. While we have a spectra representation 3 different materials
(columns of S), the constraints reduce this dimensionality to 2
 because
we only need to know the abundaces of two of the materials to
 calculate
the abundance of the third material.
%}
```

9.

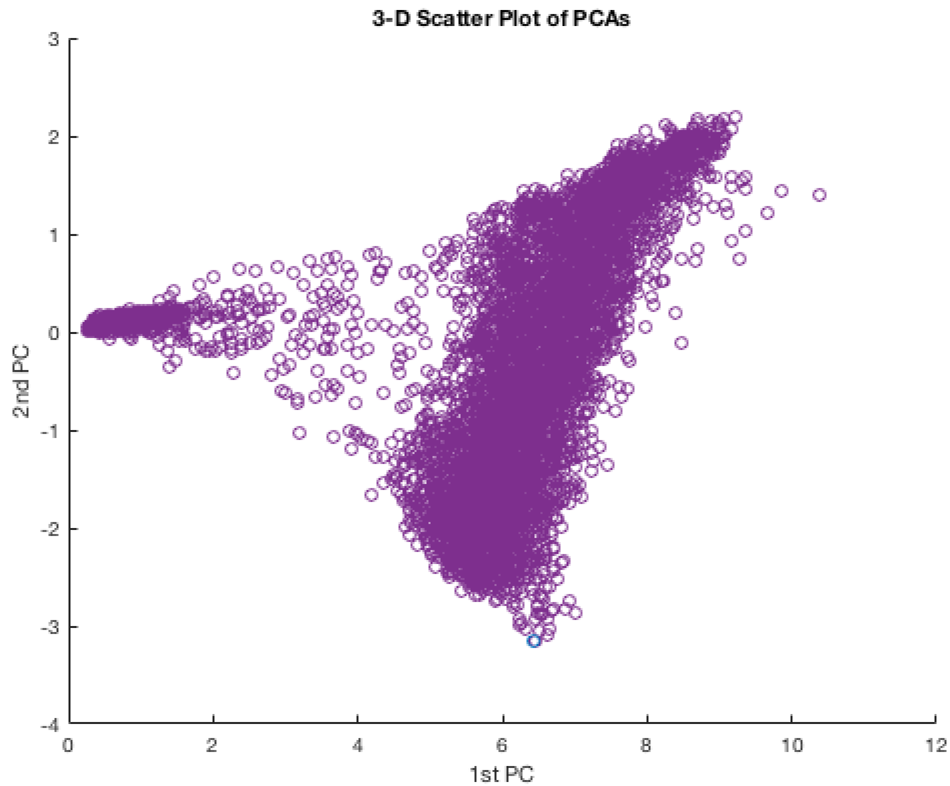
```
%{
Ux = S, where x represents the coordinates of the columns of S in the
eigenbasis
```

```
%}

coords = zeros(3,3);
for i = 1:3
    %coords of S1
    coords(1, i) = S(:,1)'*V(:,195 - i + 1);
    %coords of S2
    coords(2,i) = S(:,2)'*V(:, 195 - i + 1);
    %coords of S3
    coords(3,i) = S(:,3)'*V(:,195 - i +1);
end

figure
hold on
scatter(coords(1,1), coords(1,2),'r')
scatter(coords(2,1), coords(2,2),'r')
scatter(coords(3,1), coords(3,2),'r')
scatter(score(:, 1), score(:,2), '.')
xlabel('1st PC');
ylabel('2nd PC');
title('2-D Scatter Plot of PCAs');
hold off


figure
hold on
scatter3(coords(1,1), coords(1,2), coords(1,3))
scatter3(coords(2,1), coords(2,2),coords(2,3))
scatter3(coords(3,1), coords(3,2),coords(3,3))
scatter3(score(:, 1), score(:,2), score(:,3))
xlabel('1st PC');
ylabel('2nd PC');
zlabel('3rd PC');
title('3-D Scatter Plot of PCAs');
hold off


%{
Each column of S is located in one of the corners of our triangle
 formed
by the plot of our PCA scores. This indicates that columns of S would
 be
a good basis for our feature space because based on our discussion of
 the
constraints. The sum-to-one constraint reduces our subspace to a 2-
dimensional space and the positivy constraint reduces this space
 further
such that the boundaries represent the columns of S. Thus, we know
 that all
points will lie within a triangle looking shape where the corners
 represent
our materials. So, a plot of our points with respect to the
column space of S should be similar to our plot of the PCA scores.
```

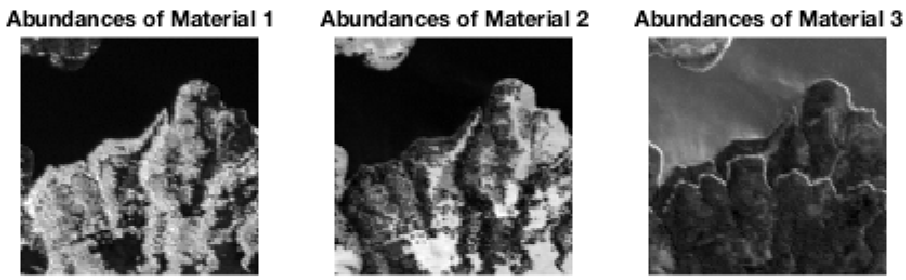
```
%}
```



2-D Scatter Plot of PCAs
2nd PC
1st PC
3
2
1
0
-1
-2
-3
-4
0
2
4
6
8
10
12




```
%10.

%Use solution to A = (S^t*S)^-1*S^t*X

temp = inv(S'*S);
A = temp*S'*img';


figure
subplot(1,3,1);
imshow(mat2gray(reshape(A(1,:)', 101, 101)))
title('Abundances of Material 1');

subplot(1,3,2);
imshow(mat2gray(reshape(A(2,:)', 101, 101)))
title('Abundances of Material 2');

subplot(1,3,3);
imshow(mat2gray(reshape(A(3,:)', 101, 101)))
title('Abundances of Material 3');
```




```
%11.
%Use linear model in 4. of theoretical part to compute A and lambda
first_right =  S'*img';
top_left = S'*S;
bottom_right = zeros();
```

Published with MATLAB® R2017a