Principal Component Analysis (PCA)

Pratheep Kumar Chelladurai

Github Link: https://github.com/pratheepkumar1/Principles-of-Color-Science/tree/main/Assignment%2014%20PCA

```
% Reading the Dataset
warning('off','all');
SGref_dataset = readtable("ccsg.xlsx");
SG_ref = SGref_dataset{:,5:end};
SGwl = (380:10:730)';
```

Question 1

Use the Matlab function **pca** to compute 36 principal components of the CCSG. Make a plot of the first 5 principal components versus wavelength, and label them with a legend. Make a plot of the cumulative percentage of variance explained as a function of number of components.

```
% Compute the principal components of 36 wavelengths using PCA function
[coeff,score,latent,tsquared,explained,mu] = pca(SG_ref);

% Displaying the principal components
coeff
```

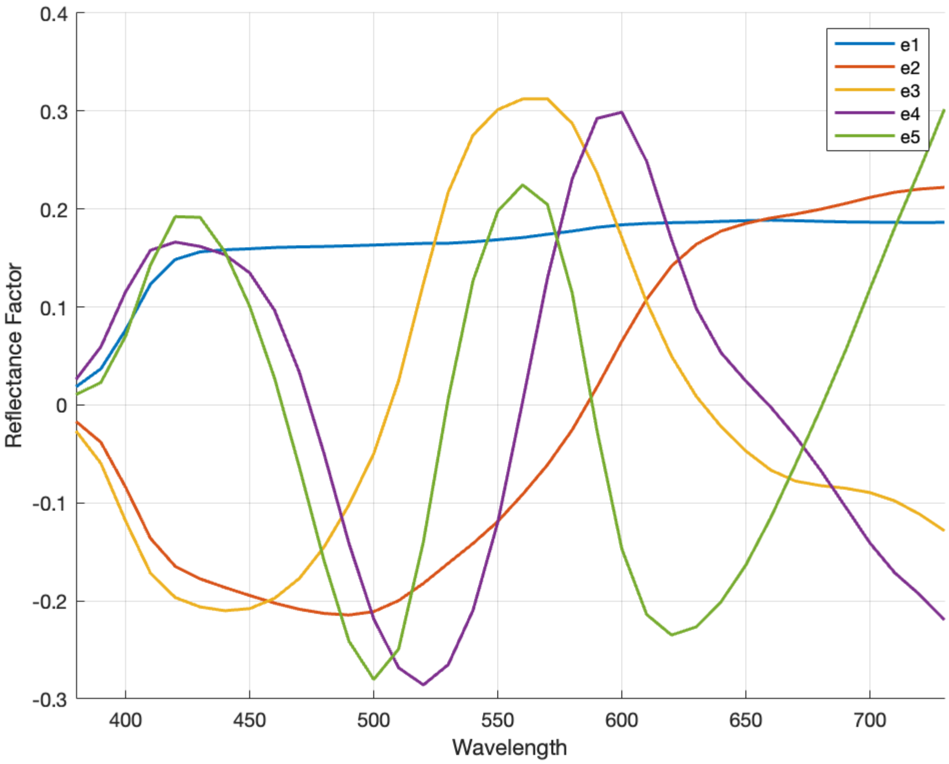
```
coeff = 36×36
    0.0186   -0.0168   -0.0265    0.0261    0.0106    0.0625    0.2049    0.0798 ···
    0.0369   -0.0381   -0.0595    0.0593    0.0229    0.1101    0.3914    0.1457
    0.0768   -0.0840   -0.1186    0.1157    0.0704    0.0942    0.5497    0.1572
    0.1233   -0.1361   -0.1717    0.1579    0.1429   -0.0257    0.4038    0.0187
    0.1485   -0.1648   -0.1966    0.1662    0.1922   -0.1106    0.1484   -0.1233
    0.1563   -0.1775   -0.2062    0.1618    0.1915   -0.1302   -0.0014   -0.1548
    0.1584   -0.1863   -0.2100    0.1534    0.1563   -0.1244   -0.1095   -0.1242
    0.1594   -0.1945   -0.2079    0.1348    0.1006   -0.1016   -0.1985   -0.0704
    0.1607   -0.2022   -0.1973    0.0966    0.0273   -0.0475   -0.2444    0.0021
    0.1613   -0.2085   -0.1770    0.0333   -0.0635    0.0393   -0.2466    0.0915
      ⋮
```

```
score
```

```
score = 140×36
    3.3336   -0.4178   -0.0054    0.0044    0.0114   -0.0210   -0.0420   -0.0081 ···
   -0.7839   -0.2294   -0.0059    0.0849   -0.0062   -0.0125    0.0419    0.0101
   -1.7847   -0.1292    0.0061    0.0545   -0.0137   -0.0170   -0.0269   -0.0053
    3.3793   -0.4512   -0.0173    0.0121    0.0145   -0.0232   -0.0323   -0.0077
   -0.7822   -0.2316   -0.0071    0.0853   -0.0056   -0.0126    0.0429    0.0104
   -1.7918   -0.1278    0.0066    0.0540   -0.0140   -0.0173   -0.0287   -0.0059
    3.3762   -0.4528   -0.0175    0.0132    0.0146   -0.0230   -0.0322   -0.0081
   -0.7816   -0.2338   -0.0082    0.0859   -0.0052   -0.0125    0.0444    0.0105
   -1.7877   -0.1290    0.0061    0.0544   -0.0138   -0.0172   -0.0276   -0.0055
    3.3635   -0.4411   -0.0127    0.0087    0.0130   -0.0230   -0.0393   -0.0089
      ⋮
```

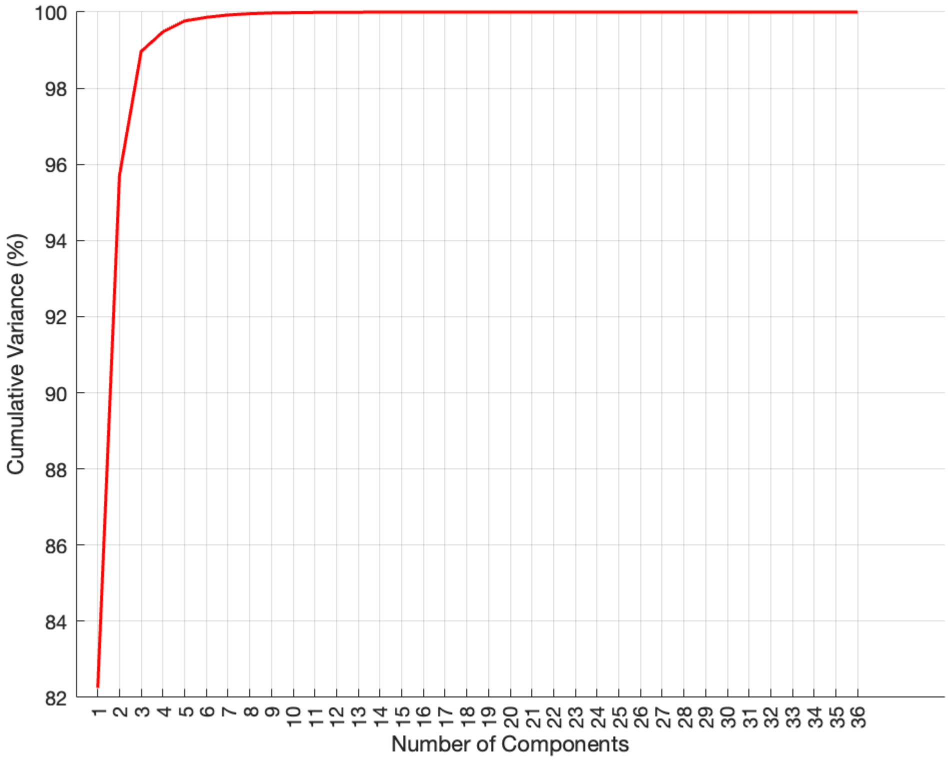
```
% Plot of the first 5 principal components versus wavelength
```

```
figure;
hold on
plot(SGwl,coeff(:,1:5),'LineWidth',1.5);
xlabel('Wavelength');
ylabel('Reflectance Factor');
xlim([380,730]);
legend("e1","e2","e3","e4","e5");
grid on
hold off
```



```
%Calculating & plotting cumulative percentage of the variance
cumsum_var = cumsum(latent);
cumsum_var_percent = (cumsum_var/cumsum_var(end,:))*100;

figure;
hold on
coeff_count = 1:1:36;
plot(coeff_count,cumsum_var_percent,'LineWidth',1.5,'Color','red');
xlabel('Number of Components');
ylabel('Cumulative Variance (%)');
xticks(1:1:36)
grid on
hold off
```



Question 2

*Use the **pca** coefficients, scores, and mu (mean) to reconstruct the 140 spectra of CCSG patches. Confirm that they match.*

```
% Reconstruct the 140 spectra of CCSG patches
SG_reconst = score*coeff'+mu
```

```
SG_reconst = 140×36
    0.1123    0.2108    0.4182    0.6675    0.8141    0.8639    0.8793    0.8872 ···
    0.0619    0.1029    0.1506    0.1747    0.1812    0.1847    0.1876    0.1893
    0.0118    0.0123    0.0121    0.0119    0.0119    0.0119    0.0118    0.0117
    0.1230    0.2220    0.4298    0.6831    0.8323    0.8821    0.8962    0.9033
    0.0628    0.1041    0.1514    0.1758    0.1823    0.1858    0.1886    0.1900
    0.0107    0.0100    0.0102    0.0103    0.0106    0.0105    0.0104    0.0103
    0.1218    0.2217    0.4302    0.6834    0.8326    0.8821    0.8961    0.9031
    0.0640    0.1049    0.1527    0.1770    0.1833    0.1864    0.1893    0.1906
    0.0115    0.0114    0.0112    0.0116    0.0115    0.0113    0.0112    0.0113
    0.1151    0.2147    0.4239    0.6775    0.8264    0.8761    0.8905    0.8983
      ⋮
```

```
%Confirm if reconstruct match
isequal(fi(SG_ref),fi(SG_reconst))
```

```
ans = logical
   1
```

Yes, the reconstructed spectral data matches with the original CCSG spectral data. For the comparison, both the matrices are converted from real numbers to fixed point data type. Because, real numbers have infinite precision and will be difficult to compare.

Question 3

*Use only the first **pca** coefficient (vector) with the mu to reconstruct the 140 CCSG patches. Compute the DE00 color difference (for 2-deg observer under D65 illuminant) between the original and reconstructed; report the mean and max DE00.*

```
% Compute CCSG reference LAB values for De00 calcualtion
[num,txt] = xlsread('all_1nm_data.xlsx');
cmf2 = interp1(num(:,1),num(:,6:8),SGwl);
D65 = interp1(num(:,1),num(:,3),SGwl);
SGXYZ_ref = ((cmf2' * diag(D65) * SG_ref') ./ (cmf2(:,2)' * D65) )';
SGLab_ref = xyz2lab(SGXYZ_ref);

% Reconstructed CCSG data with 1st PCA coefficient
SG_reconst_first = score(:,1)*coeff(:,1)'+mu
```

```
SG_reconst_first = 140×36
    0.1171    0.2169    0.4132    0.6259    0.7373    0.7737    0.7865    0.7946 ···
    0.0407    0.0649    0.0970    0.1183    0.1260    0.1303    0.1344    0.1382
    0.0221    0.0280    0.0202   -0.0051   -0.0226   -0.0261   -0.0241   -0.0213
    0.1180    0.2186    0.4167    0.6316    0.7441    0.7809    0.7938    0.8019
    0.0407    0.0650    0.0972    0.1185    0.1262    0.1306    0.1347    0.1385
    0.0220    0.0277    0.0197   -0.0060   -0.0237   -0.0272   -0.0252   -0.0225
    0.1179    0.2185    0.4165    0.6312    0.7437    0.7804    0.7933    0.8014
    0.0407    0.0650    0.0972    0.1186    0.1263    0.1307    0.1348    0.1386
    0.0221    0.0278    0.0200   -0.0055   -0.0231   -0.0266   -0.0245   -0.0218
    0.1177    0.2180    0.4155    0.6296    0.7418    0.7784    0.7913    0.7993
     ⋮
```

```
% Compute CCSG 1st PCA coefficient data's LAB values
SGXYZ_reconst_first = ((cmf2' * diag(D65) * SG_reconst_first') ./ ...
    (cmf2(:,2)' * D65) )';
SGLab_reconst_first = xyz2lab(SGXYZ_reconst_first);

% DE00 color difference between the original and reconstructed
De00_first = deltaE00(SGLab_reconst_first',SGLab_ref');
```

```
% De00 mean
De00_first_mean = mean(De00_first)
```

```
De00_first_mean = 18.4574
```

```
% De00 Max Value
De00_first_max = max(De00_first)
```

```
De00_first_max = 39.5881
```

Question 4

*Reconstruct the 140 CCSG patches with the first two **pca** coefficients, and report mean and max DE00.*

```
% Reconstructed CCSG data with first two PCA coefficient
SG_reconst_second = score(:,1:2)*coeff(:,1:2)'+mu
```

```
SG_reconst_second = 140×36
    0.1241    0.2328    0.4483    0.6828    0.8062    0.8479    0.8643    0.8758 ···
    0.0446    0.0736    0.1163    0.1495    0.1638    0.1710    0.1772    0.1828
    0.0243    0.0329    0.0311    0.0125   -0.0013   -0.0032    0.0000    0.0038
    0.1256    0.2358    0.4546    0.6930    0.8185    0.8610    0.8778    0.8896
    0.0446    0.0738    0.1166    0.1500    0.1644    0.1717    0.1779    0.1835
    0.0241    0.0326    0.0304    0.0114   -0.0026   -0.0045   -0.0014    0.0024
    0.1255    0.2357    0.4545    0.6928    0.8183    0.8608    0.8776    0.8894
    0.0447    0.0739    0.1169    0.1504    0.1648    0.1721    0.1784    0.1840
    0.0242    0.0328    0.0308    0.0121   -0.0018   -0.0037   -0.0005    0.0033
    0.1251    0.2348    0.4526    0.6897    0.8145    0.8567    0.8734    0.8851
      ⋮
```

```
% Compute CCSG 2nd PCA coefficient data's LAB values
SGXYZ_reconst_second = ((cmf2' * diag(D65) * SG_reconst_second') ./ ...
    (cmf2(:,2)' * D65) )';
SGLab_reconst_second = xyz2lab(SGXYZ_reconst_second);

% DE00 color difference between the original and reconstructed
De00_second = deltaE00(SGLab_reconst_second',SGLab_ref');
```

```
% De00 mean
De00_second_mean = mean(De00_second)
```

```
De00_second_mean = 14.3649
```

```
% De00 Max Value
De00_second_max = max(De00_second)
```

```
De00_second_max = 44.1072
```

Bonus

*Keep going with this series, and determine how many **pca** coefficients are necessary to reduce the max DE00 below 0.5.*

```
% Determine number of pca coefficients to reduce the max DE00 below 0.5.
for i = 1:36
    SG_reconst_f = score(:,1:i)*coeff(:,1:i)'+mu;
    SGXYZ_reconst_f = ((cmf2' * diag(D65) * SG_reconst_f') ./ ...
        (cmf2(:,2)' * D65) )';
    SGLab_reconst_f = xyz2lab(SGXYZ_reconst_f);
    De00_f = deltaE00(SGLab_reconst_f',SGLab_ref');
    De00_f_max = max(De00_f);
    if(De00_f_max <= 0.5)
        disp("Number of coefficients to minimize De00 below 0.5 is " + i)
        disp("The De00 value at " + i + "th coefficient is " + De00_f_max);
        break;
    end
end
```

```
Number of coefficients to minimize De00 below 0.5 is 11
The De00 value at 11th coefficient is 0.34552
```

Functions

```
function   De00=deltaE00(Lab1, Lab2)

%CIELAB Chroma
C1 = sqrt(Lab1(2,:).^2+Lab1(3,:).^2);
C2 = sqrt(Lab2(2,:).^2+Lab2(3,:).^2);

%Lab Prime
mC = (C1+C2)./2;
G=0.5*(1-sqrt((mC.^7)./((mC.^7)+(25.^7))));
LabP1 = [Lab1(1,:) ; Lab1(2,:).*(1+G) ; Lab1(3,:)];
LabP2 = [Lab2(1,:) ; Lab2(2,:).*(1+G) ; Lab2(3,:)];

%Chroma
CP1 = sqrt(LabP1(2,:).^2+LabP1(3,:).^2);
CP2 = sqrt(LabP2(2,:).^2+LabP2(3,:).^2);

%Hue Angle
hP1t = atan2Deg(LabP1(3,:),LabP1(2,:));
hP2t = atan2Deg(LabP2(3,:),LabP2(2,:));

%Add in 360 to the smaller hue angle if absolute value of difference is > 180
hP1 = hP1t + ((hP1t<hP2t)&(abs(hP1t-hP2t)>180)).*360;
hP2 = hP2t + ((hP1t>hP2t)&(abs(hP1t-hP2t)>180)).*360;

%Delta Values
DLP = LabP1(1,:) - LabP2(1,:);
DCP = CP1 - CP2;
DhP = hP1 - hP2;
DHP = 2*(CP1.*CP2).^(1/2).*sinDeg(DhP./2);

%Arithmetic mean of LCh' values
mLP = (LabP1(1,:)+LabP2(1,:))./2;
mCP = (CP1+CP2)./2;
mhP = (hP1+hP2)./2;
```

```matlab
%Weighting Functions
SL = 1+(0.015.*(mLP-50).^2)./sqrt(20+(mLP-50).^2);
SC = 1+0.045.*mCP;
T = 1-0.17.*cosDeg(mhP-30)+0.24.*cosDeg(2.*mhP)+0.32.* ...
    cosDeg(3.*mhP+6)-0.2.*cosDeg(4.*mhP-63);
SH = 1+0.015.*mCP.*T;

%Rotation function
RC = 2.*sqrt((mCP.^7)./((mCP.^7)+25.^7));
DTheta = 30.*exp(-((mhP-275)./25).^2);
RT = -sinDeg(2.*DTheta).*RC;

%Parametric factors
kL = 1;
kC = 1;
kH = 1;

De00 = ((DLP./kL./SL).^2+(DCP./kC./SC).^2+(DHP./kH./SH).^2+...
    (RT.*(DCP./kC./SC).*(DHP./kH./SH))).^(1/2);

end

%---------------
function out = sinDeg(in);
out = sin(in.*pi./180);
end

%---------------
function out = cosDeg(in);
out = cos(in.*pi./180);
end

%---------------
function out = atan2Deg(inY,inX);
out = atan2(inY,inX).*180./pi;
out = out+(out<0).*360;

end
```