

Simulation study for interpretation of PCA loading spectra

Takuma Genkawa, Akifumi Ikehata

(Institute of Food Research, NARO, Tsukuba, Japan, 305-8642 E-mail: genkawat748@affrc.go.jp)



Introduction

Principal component analysis (PCA) is often used in analyzing NIR spectra. The NIR spectra are expressed in terms of score, loading, and contribution ratio, the score indicates quantitative information, the loading indicates spectroscopic information, and the contribution ratio indicates the importance of each PCs.

The scores are used to construct calibration models and to draw spectral images, and the loadings are interpreted to know the scientific basis of the models and images. In the present study, to help interpretation of loading spectra, simulated model spectra comprising Lorentzian peaks were subjected to PCA and obtained loading spectra were examined.

Model Data and Simulation

Figure 1 shows two source spectra, A and B, which will be used as the pure component spectra. The source spectra A and B were comprising Lorentzian peaks and centered at channels 40 and 60. Using these source spectra, four type of model spectra are prepared by referring simulation study of 2DCOS¹ as shown Fig. 2.

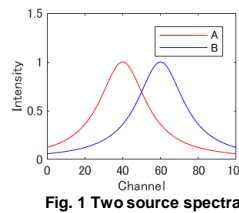
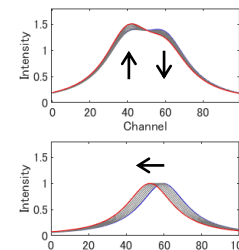


Fig. 1 Two source spectra

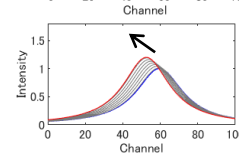
1) Two separate bands with intensity changes

The two source spectra are linearly combined while intensity changes for the spectra A and B have opposite directions.



2) Band position shift

A single band of spectra B shifts with a fixed intensity and line shape.



3) Band position shift coupled with intensity change

A single band of spectra B shifts coupled with intensity increase.

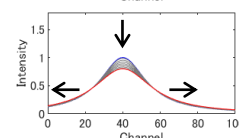


Fig. 2 Model spectra

4) Line broadening

The band width of spectra A is increased with decreasing its height while the integral intensity of spectra A is kept at a constant value.

Results and Discussion

Results of PCA (Figures 4-6) are displayed with analyzed spectra (Fig. 3a), mean spectrum (Fig. 3b), mean-centered spectra (Fig. 3c), explained X-variance (spectral variance, Fig. 3d), loading spectra of PC1 to PC4 (Fig. 3e), scores of PC1 to PC4 (Fig. 3f), and reconstructed spectra with PC1 to PC4 (Fig. 3g). The reconstructed spectra are calculated using the loading spectrum and score of each PC.

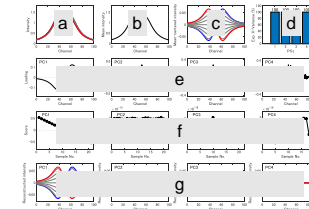


Fig. 3 Layout of PCA results

Spectra with two overlapped bands changing in intensity in opposite directions give the PC1 loading spectrum with large values in both positive (channel 60) and negative (channel 40) side and the PC1 score decreasing linearly. Because the intensity changes simultaneously, the explained X-variance reaches 100%, so that the mean-centered spectra are equal to the reconstructed spectra calculated from the PC1 score and loading spectrum.

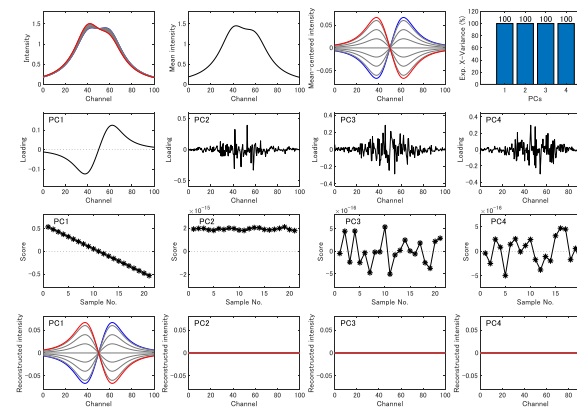


Fig. 4 PCA results of model spectra with two overlapped bands changing in intensity in opposite directions

Spectra with a single band shifting also give the PC1 loading spectrum with large positive and negative values and the PC score increasing linearly (Fig. 5). The mean spectrum has a single peak, and the explained X-variance of PC1 is 98.8%. The PC2 loading spectrum looks like the second derivative of the mean spectrum, and the PC2 score increases from the first sample (No. 0) to the middle one (No. 11), then decreased. The reconstructed spectra of PC2 have small but important spectral changes required to describe the band shifting.

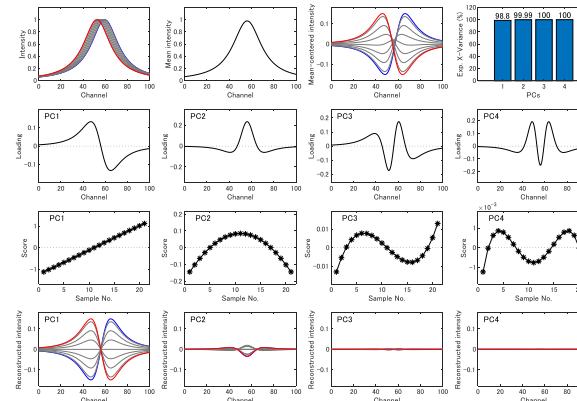


Fig. 5 PCA results of model spectra with a single band shift

As shown in Fig. 6, the PC1 loading spectrum of spectra with a band simultaneously shifting in position and increasing in intensity is similar to that of a band shift (Fig. 5) but not positive and negative symmetrical. The PC2 loading spectrum shows different shape from that of a band shift. Asymmetry of loading spectrum seems to be a good marker for concentration changes that occur simultaneously with a band shift, which is often observed for spectra of solution mixtures.¹

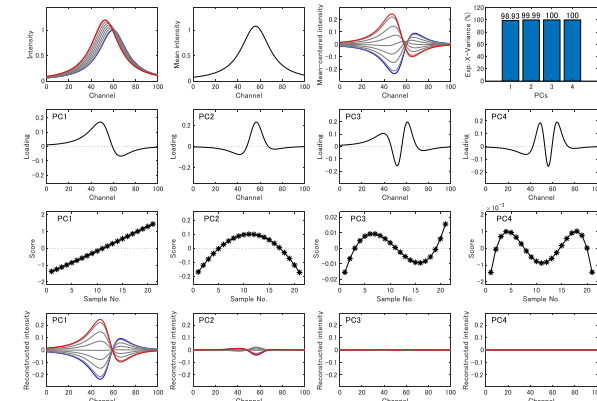


Fig. 6 PCA results of model spectra with a single band simultaneously shifting in position and increasing in intensity

Spectra in which a band broadening give the PC1 loading spectrum with large positive values and the PC score decreasing linearly (Fig. 7). Small negative loading values that might be overlooked are obtained on both side of the peak. The reconstructed spectra of PC1 enhanced spectral changes corresponding to these small negative loading values, as same as 2D correlation spectra in which four-way symmetric pattern appears.¹

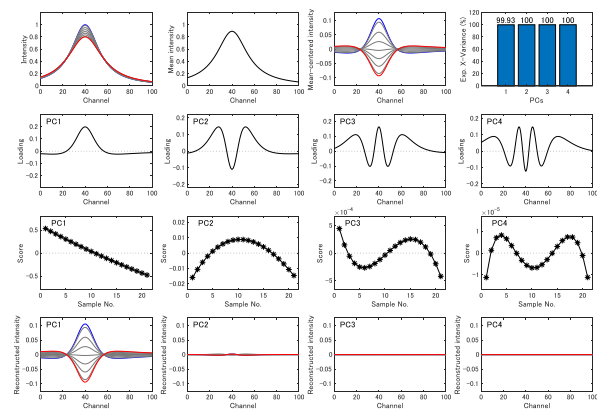


Fig. 7 PCA results of model spectra with band broadening

Conclusion

For deeper interpretation of loading spectra, PCA results of four-type model spectra were compared. The score showed linear change regardless the type of spectra, and loading spectra are also similar to each other except for spectra with a band broadening. Symmetry of loading spectrum and spectral variation in reconstructed spectra will give an insight into chemical information related to the PCA subjected spectra.

Reference

- 1) Noda I. and Ozaki Y., Two-dimensional correlation spectroscopy, Wiley, 56-64, 2004