Getting second order advantage for first order data

A. Naseri

University of Tabriz
E-mail: ab.naseri@gmail.com
Definition of second order advantage

Determination of anlayte(s) in the presence of the unexpected interferent(s) in the prediction samples.

Second-order calibrations can predict the concentration of the analyte of interest even in the presence of unexpected interferents, this property have been named second-order advantage.

Analyte’s determinations in unknown matrix or second order advantage? Which one?


H. Abdollahi, M. Kooshki, Electroanalysis 2010, 22, No. 19, 2245 – 2253
First order calibration methods can not predict the concentration of analyte in presence of unexpected interferent.

Is it correct in any condition?
H-point standard addition method

Main disadvantage of HPSAM: interference spectrum must be known.
Generalized H-point standard addition method

*Development of H-point standard additions method for analyte determinations in unknown matrix. Location of linear matrix spectral interval*


*Generalized H-point standard addition method for analyte determinations in unknown matrix*


This method was presented for covering disadvantages of HPSAM
The first requirement is that the unknown interferent has to show a linear behavior between $\lambda_1$ and $\lambda_2$.

The second restriction for the selection of these wavelengths is related to the analyte, which has to show absorbance values that are not linearly related in the wavelength interval $[\lambda_1, \lambda_2]$. 

X: analyte  Y: interference  S: sample
Steps of method:

Selecting of three wavelengths in this interval ($\lambda_1$, $\lambda_m$ and $\lambda_k$).

Interference shows a linear behavior between $\lambda_1$ and $\lambda_k$:

\[ A_{Y,j} = A_{Y,1} + b \lambda_j \]

\[ A_{s,j} = A_{X,j} + A_{Y,j} = C_X \varepsilon_{X,j} + A_{Y,1} + b \lambda_j \]

In order to simplify, $\lambda_m$ select as the average wavelength between $\lambda_1$ and $\lambda_k$. 
Steps of method

*Standard addition for analyte and record of absorbance at three selected wavelength.*

We can define two new parameters, as the difference between $A_m$ and $A_1$, $A_k$ and $A_m$:

$$
\Delta A_1 = A_m - A_1 = \left( A_{X,m}^0 - A_{X,1}^0 \right) + (A_{Y,m} - A_{Y,1}) + (\varepsilon_{X,m} - \varepsilon_{X,1})C_i
$$

$$
\Delta A_m = A_k - A_m = \left( A_{X,k}^0 - A_{X,m}^0 \right) + (A_{Y,k} - A_{Y,m}) + (\varepsilon_{X,k} - \varepsilon_{X,m})C_i
$$

\[
for \lambda_1: \quad A_1 = A_{X,1}^0 + A_{Y,1} + \varepsilon_{x,1} C_i \\
for \lambda_m: \quad A_m = A_{X,m}^0 + A_{Y,m} + \varepsilon_{x,m} C_i \\
for \lambda_k: \quad A_k = A_{X,k}^0 + A_{Y,k} + \varepsilon_{x,k} C_i
\]
From the abscissa of their intersection point, which is H-point \((-C_H, \Delta A_H)\)

\[
A_m - A_1 = A_K - A_m
\]

\[
(A^0_{x,m} - A^0_{x,1}) + (A^0_{y,m} - A^0_{y,1}) + (\varepsilon_{x,m} - \varepsilon_{x,1})C_i = (A^0_{x,m} - A^0_{x,1}) + (A^0_{y,m} - A^0_{y,1}) + (\varepsilon_{x,m} - \varepsilon_{x,1})C_i
\]

\[
(A^0_{y,m} - A^0_{y,1}) = (A^0_{y,m} - A^0_{y,1})
\]

\[
-C_H = \frac{(2A^0_{x,m} - A^0_{x,1} - A^0_{x,k})}{(\varepsilon_1 + \varepsilon_k - 2\varepsilon_m)}
\]
The main question

How can we locate the linear spectral intervals for the unknown matrix?

Two approaches have been described to locate the linear spectral behavior for the unknown interferent.

- If the interferent shows a linear behavior in the wavelength interval considered, the second derivative of the sample absorbance with respect to the wavelength should only contain a contribution of the analyte.

\[
\frac{d^2 A_s}{d\lambda^2} = \frac{d^2 A_x^0}{d\lambda^2} + \frac{d^2 A_y}{d\lambda^2} = \frac{d^2 A_x^0}{d\lambda^2}
\]

Therefore, the linear spectral range of the interferent can be found as the region with a constant ratio of the second derivative of the sample and reference.

\[
\text{Ratio} = \frac{\frac{d^2 A_s}{d\lambda^2}}{\frac{d^2 A_x^{\text{ref}}}{d\lambda^2}} = \frac{C_X^0}{C_X^{\text{ref}}}
\]
A plot of Ratio versus $\lambda_j$ allows to locate the wavelength interval in which the ratio is constant and consequently the interferent shows a linear behavior.
Second approach to locate the linear spectral behaviour of the unknown interferent

If the spectral behavior of the unknown interferent \( Y \) in the wavelength interval is linear,

So the first derivative of \( A_Y \) will be constant:

\[
A'_{Y,j} = A_{Y,1} + b \lambda_j
\]

\[
A'_Y = b
\]

A plot of the values of the difference between the first derivative of the sample and the first derivative of the reference analyte solution versus the first derivative of the reference solution results in a straight line with intercept \( b \) and slope \[
\frac{C^0_x - C^\text{ref}_x}{C^\text{ref}_x} \]

\[
A'_s - A'_x^{\text{ref}} = b + (C^0_x - C^\text{ref}_x) \varepsilon'_x = b + \frac{(C^0_x - C^\text{ref}_x)}{C^\text{ref}_x} A'_x^{\text{ref}}
\]
% Generalized N-point Standard Addition Method

clear

close all

% construction of spectra

x=[350:1:550];

for i=1:2;
    for j=1:length(x);
        A=[1.5 1];
        M=[380 420];
        S=[12 16];
        spect(j,i)=A(1,i)*exp(\{-(x(j)-M(1,i)).^2\}/(2*S(1,i)^2));
    end
end

Y=spect(:,2);

aa=input('do you want the spectra of interferent be linear ? (y/n)', 's');

if aa=='y'
    Y=0.14+Y*0.0008;Y=Y;
end

figure(1),plot(x,Y), xlabel('Wavelength(um)', 'FontSize',14), ylabel('A', 'FontSize',14), title('Spectra')

pause

C_un=input('pls enter Conc. of analyte in sample i.e.[0.3]');

X=C_un*spect(:,1);

hold on

figure(1),plot(x,X,'r')

hold on

pause

S=[C_un*spect(:,1)+Y];

figure(1),plot(x,S,'g'),

figure(1),plot(x,S,'g'),

for i=1:(length(S)-1)
    deriv_S(i)=(S(i+1)-S(i))/x(i+1)-x(i(i);
    deriv_Y(i)=(Y(i+1)-Y(i))/x(i+1)-x(i(i);
    end
do you want the spectra of interferent be linear? (y/n) y
do you want the spectra of interferent be linear? (y/n) y
please enter Conc. of analyte in sample i.e. [0.3] [0.6]

C_un =

0.6000
do you want the spectra of interfering be linearly
plz enter Conc. of analyte in sample i.e. [0.3][0.5]

C_un =

0.6000

plz enter Conc. of standard added of analyte i.e.

C_added =

0.2000 0.4000 0.6000

Wavelength(nm)
do you want the spectra of interferent be linear? (y/n) y
plz enter Conc. of analyte in sample i.e. [0.3] [0.6]

C_un =

0.6000

plz enter Conc. of standard added of analyte i.e. [0.1 0.2 0.4] [0]

C_added =

0.2000 0.4000 0.6000

plz enter three selected wavelength i.e. [392 357 402] [350 355 405]

C_pre =

0.6000

>>
do you want the spectra of interferent be linearly
plz enter Conc. of analyte in sample i.e.[0.3][0.2]

\[ C_{un} = 0.5000 \]
do you want the spectra of interferent be linear?
plz enter Conc. of analyte in sample i.e. [0.3][0.4]

C_un =

0.5000

plz enter Conc. of standard added of analyte in

C_added =

0.2000  0.4000  0.6000
do you want the spectra of interferent be linear?

plz enter Conc. of analyte in sample i.e. [0.3] [1.0]

\[ C_{un} = \]
\[ 0.5000 \]

plz enter Conc. of standard added of analyte i.e. [0.2000] [0.4000] [0.6000]

\[ C_{added} = \]
\[ 0.2000 \quad 0.4000 \quad 0.6000 \]

plz enter three selected wavelength i.e. [392 397 402]

\[ C_{pre} = \]
\[ 0.3915 \]
H-point curve isolation

Curve resolution procedure for isolating the spectra of unknown interferences from the sample spectrum in analyte determination
The method is based on the selection of a reference wavelength and then defining $K_i$ as:

$$K_i = X_{\text{ref}} / X_i$$

Where $X_{\text{ref}}$ is the absorbance of the analyte at reference wavelength $K_i$ calculated from spectrum of pure compound.
\[ S_{\text{ref}} - K_i S_i = X_{\text{ref}} + Y_{\text{ref}} - K_i (X_i + Y_i) \]

Rearrangement gives:

\[ S_{\text{ref}} - K_i S_i = Y_{\text{ref}} - K_i Y_i \]

Hence we have a set of \( n \) equations:

\[
\begin{align*}
S_{\text{ref}} - K_1 S_1 &= Y_{\text{ref}} - K_1 Y_1 \\
S_{\text{ref}} - K_2 S_2 &= Y_{\text{ref}} - K_2 Y_2 \\
S_{\text{ref}} - K_3 S_3 &= Y_{\text{ref}} - K_3 Y_3 \\
&\vdots \\
S_{\text{ref}} - K_n S_n &= Y_{\text{ref}} - K_n Y_n
\end{align*}
\]

\[
\begin{align*}
Y_1 &= \left( (S_{\text{ref}} - K_1 S_1) - Y_{\text{ref}} \right) / -K_1 \\
Y_2 &= \left( (S_{\text{ref}} - K_2 S_2) - Y_{\text{ref}} \right) / -K_2 \\
Y_3 &= \left( (S_{\text{ref}} - K_3 S_3) - Y_{\text{ref}} \right) / -K_3 \\
&\vdots \\
Y_n &= \left( (S_{\text{ref}} - K_n S_n) - Y_{\text{ref}} \right) / -K_n
\end{align*}
\]

Every things in above equations are known except \( Y_{\text{ref}} \).
The main problem:

How can we estimate the absorbance of the interferent at the reference wavelength ($Y_{\text{ref}}$)?

- The maximum value of the $Y_{\text{ref}}$ ($Y_{\text{ref}}^{\text{max}}$) cannot be higher than $S_{\text{ref}}$.
- The minimum value of the $Y_{\text{ref}}$ is the smallest absorbance value that provides non-negative absorbance values for the interferent spectrum at any wavelength calculated from above equations.
We can plot several interferent spectra from above equations considering the range described by the minimum and maximum absorbance values previously selected and varying $Y_{\text{ref}}$ within this range to produce meaningful spectral differences from one spectrum to the next.

The real unknown interferent spectrum will be one of all calculated spectra.
• To calculate the $Y_{ref}$, it is necessary to find a pair of wavelength with the same absorbance value in every hypothetical interferent spectrum although different in magnitude from one spectrum to the other.

Wavelength $m$ & $n$ represent one of several pairs of wavelength at which

$$y_m = y_n$$
We can write:

\[ S_{\text{ref}} - K_m S_m = Y_{\text{ref}} - K_m Y_m \]

\[ S_{\text{ref}} - K_n S_n = Y_{\text{ref}} - K_n Y_n \]

\[ Y_m = Y_n \]

We have three equations and three unknowns, so we can calculate the value of \( Y_{\text{ref}} \) from:

\[ Y_{\text{ref}} = S_{\text{ref}} - K_m S_m + \left\{ K_m \left[ (S_{\text{ref}} - K_m S_m) - (S_{\text{ref}} - K_n S_n) \right] / (K_n - K_m) \right\} \]
It is possible to estimate the unknown spectrum.

\[ Y_1=\frac{(S_{\text{ref}}-K_1 S_1)-Y_{\text{ref}}}{-K_1} \]
\[ Y_2=\frac{(S_{\text{ref}}-K_2 S_2)-Y_{\text{ref}}}{-K_2} \]
\[ Y_3=\frac{(S_{\text{ref}}-K_3 S_3)-Y_{\text{ref}}}{-K_3} \]
\[ \vdots \]
\[ \vdots \]
\[ \vdots \]
\[ Y_n=\frac{(S_{\text{ref}}-K_n S_n)-Y_{\text{ref}}}{-K_n} \]
Criteria for selecting of reference wavelength:

All measured wavelength are assayed as the reference wavelength but wavelength is select that provide:

• Greater number of $S_{\text{ref}} - K_i S_i$

• Minimum standard deviation of $S_{\text{ref}} - K_i S_i$ (n=number of processed samples)
% H-point Curve Isolation Method
clc
clear
close all
% construction of spectra
x=[350:1:550];
for i=1:4;
    lambda=[0.8 0.6 0.5 0.7];
    M=[380 460 430 480];
    S=[20 30 25 31];
    spect(:,1)=lambda(1,1)*exp((-((x(j)-M(1,1)).^2)/(2*S(1,1)^2)));
end
SpectA=[spect(:,1)+spect(:,2)];
figure(1),plot(x,SpectA,'b'),xlabel('Wavelength (nm)', 'FontSize', 14), ylabel('A', 'FontSize', 14), title('Spectra')
SpectI=[spect(:,3)+spect(:,4)];
hold on
figure(1),plot(x,SpectI,'g')
SpectS=[SpectA*SpectI];
noise=rand(size(SpectS))*0.01;
% Spect_S=Spect_S+noise;
hold on
figure(1),plot(x,SpectS,'r')
% selection of reference wavelength
Ref_wav=input('plz enter Reference wavelength i.e. [470]');
[p,q]=find(x==Ref_wav);
K=(SpectA(p,q)*ones(size(x)))./SpectA;
figure(2),plot(x,K,'*')
% prediction of the spectrum of interferent
I_ref=SpectS(p,q).*0.02;0.2;
for i=1:length(x)
    for j=1:length(I_ref)
        err(i,j)=abs(I_ref(j,i)/(spect(i,j)-spect(i,j)));
    end
end
The selected range for $Y_{ref}$ must be limited.

Two selected wavelength are 450 and 471 nm.
The pair wavelengths are not selected correctly.
After extraction of the spectrum of interferent, it is possible to determine analyte correctly using methods such as HPSAM.
Different procedures have been devised in order to quantitate analytes from first-order multivariate data in the presence of unexpected substances.

One alternative is to select a spectral window where the contribution from the interferences is relatively low in comparison to the analyte contribution. This is for instance possible by resorting to the so-called net analyte signal regression plots.
The net analyte signal is the part of the total signal which can be uniquely ascribed to the analyte of interest, and the regression plot corresponds to the net analyte signal for a test sample as a function of that for the pure analyte.

This plot should ideally be a straight line. Significant deviations from linearity indicate the presence of unexpected constituents, opening the possibility of selecting only those sensor regions where these plots are linear.
Augmentation of first-order instrumental data

The spectral data matrix $D$ contains in its rows the individual spectra measured for the different analyzed samples, and in its columns the sample signals measured at each spectral wavelength.

In this way, calibration can be carried out with few standards instead of a large set of calibration standards and, in addition, it is possible to quantify the analytes in the presence of potential interferences.
Flow chart of MCR-ALS

\[ D = CS^T + E \] (bilinear model)

Data matrix decomposition according to a bilinear model

\[ \min_{\hat{C}, \text{constraints}} \| \hat{D}_{\text{PCA}} - \hat{C} \hat{S}^T \| \]

\[ \min_{S^T, \text{constraints}} \| \hat{D}_{\text{PCA}} - \hat{C} \hat{S}^T \| \]

Resolved Spectra profiles

Resolved Concentration profiles

Results of the ALS optimization procedure: Fit and Diagnostics

From the home page of R. Tauler
Rank estimation

The number of chemical species present in augmented matrix was first estimated by singular value decomposition (SVD).

This initial number of components can be afterwards refined considering larger or lower number of components, and checking for their fit and reliability.
Initial estimates

Concentration or spectral profiles?

Initial estimates of the pure spectra were obtained by using the SIMPLISMA algorithm, a technique based on selecting of the purest variables. SIMPLISMA was applied to the augmented matrix to search for spectral estimates of the different components.
Constraints

Non-negativity

Non-negativity constraints are applied to the concentration profiles, due to the fact that the concentrations of the chemical species are always positive values or zero. Non-negativity constraints are also applied for UV-Vis, fluorescence or near-infrared spectra.

equality constraint

The applied equality constraint, which consists of fixing the pure spectra of the Analytes during the optimization process, contributed to break-up the rotational and intensity ambiguities that are intrinsic to the MCR-ALS methods.
How we can implement the known values of concentrations?
Thanks for your attention