

Information recovery from near infrared data

R. S. Anderssen¹F. R. de Hoog²I. J. Wesley³

(Received 22 January 2011; revised 7 June 2011)

Abstract

In many practical situations, classical modelling protocols are either inappropriate or infeasible for the recovery of an estimate of some specific property of an object, such as the protein content of wheat, from available indirect (spectroscopic) measurements. In such situations, some form of calibration-and-prediction (machine learning) is a popular alternative. For a representative set of samples, inexpensive and rapidly available indirect (spectroscopic) encapsulations of the property are recorded for each sample along with an independent laboratory measurement of the value of the specified property. Because many more indirect measurement values are recorded for each sample than the number of samples tested, the resulting system is highly under-determined in the sense of performing the calibration step: the identification of a predictor which can be applied to the indirect measurements of a new sample to predict its value of the property. Various dimension reduction methodologies have been proposed for performing the calibration step, including principal component regression, partial least squares, independent component analysis and neural network analysis. Independently, because of the high accuracy with

<http://anziamj.austms.org.au/ojs/index.php/ANZIAMJ/article/view/3909>

gives this article, © Austral. Mathematical Soc. 2011. Published July 10, 2011. ISSN 1446-8735. (Print two pages per sheet of paper.) Copies of this article must not be made otherwise available on the internet; instead link directly to this URL for this article.

which near infrared spectra are recorded using computer controlled instrumentation, derivative spectroscopy techniques can be utilised to explore differences in the molecular structure of cereal grains. For the optimisation of the recovery of estimates of the property from the indirect measurements of new samples, two questions are explored in this article; one practical, the other theoretical: (i) To what extent should preprocessing (such as fourth differentiation) be applied to the indirect measurements before the calibration step is performed? (ii) Is there an algorithmic advantage in viewing partial least squares as an implementation of simultaneous minimisation?

Contents

1	Introduction	C335
2	Introductory background	C337
2.1	NIR spectroscopy	C338
2.2	Calibration and prediction methodologies	C338
2.3	Derivative spectroscopy	C339
3	Preprocessing of the spectra	C339
3.1	The fourth differentiation possibility	C340
3.1.1	Theoretical considerations	C340
3.1.2	SVD analysis of wheat NIR data	C341
4	Simultaneous minimisation	C344
	References	C346

1 Introduction

When the complexity of a situation is such that it is not possible and/or appropriate to formulate a mathematical model that relates the inputs to the outputs, a popular alternative is to perform the modelling using some form of calibration-and-prediction (CAP), introduced in section 2. The overall goal is the determination of a predictor for some specified property \mathcal{P} of an object, such as the protein content of wheat, where the predictor is defined in terms of some indirect (spectroscopic) measurements that encapsulate the essence of \mathcal{P} . Calibration is used to estimate the structure of the predictor.

For a set of m representative samples, two independent classes of measurements are performed:

1. For each sample, a large comprehensive set of inexpensive and rapidly recorded indirect measurements are collected. A popular choice is a near-infrared (NIR) spectrum over a comprehensive range of wavelengths (section 2). Depending on the circumstances, other forms, such as Raman spectra [2] and grain hardness measurements [4], are utilised either independently of or jointly with the NIR spectra. The overall rationale motivating the choice of the indirect measurements is that (localised) features in them are strongly linked to the information required about \mathcal{P} .
2. For each sample, a (possibly expensive, time consuming and/or hazardous) laboratory measurement is performed to determine the value of \mathcal{P} (e.g., protein content).

This set of dual measurements defines the ‘calibration data’. In the calibration step, a (linear) relationship, that defines the property as a function of the indirect measurements, is fitted to the calibration data. For NIR measurements, where a linear relationship is justified (Beer’s Law) [14], the resulting algebraic equations take the form

$$\mathbf{S}\boldsymbol{\beta} = \mathbf{p}, \quad (1)$$

where the $(m \times n)$ -matrix \mathbf{S} , with $n \gg m$, has the m measured spectral responses for n evenly spaced wavelengths stored as its rows, \mathbf{p} is the vector of the measured values of the property \mathcal{P} , and $\boldsymbol{\beta}$ is the unknown (regression) coefficient vector to be determined. It is this linear NIR situation which will be the focus of this article.

Since only a subset of the spectral responses are related to the property \mathcal{P} [14], the estimation of $\boldsymbol{\beta}$ as $\hat{\boldsymbol{\beta}}_k$ will involve a dimension reduction calibration step resulting in the replacement of equation (1) by

$$\mathbf{S}_k \hat{\boldsymbol{\beta}}_k = \mathbf{p}, \quad \mathbf{S}_k \in \mathbb{R}^{m \times n}, \quad (2)$$

with $\text{rank}(\mathbf{S}_k) = k < m$. It is the $\hat{\boldsymbol{\beta}}_k$ estimate, generated at the calibrations step, that defines the predictor \mathbf{p}^* of \mathcal{P} for a new sample with NIR spectrum \mathbf{s}^* ; namely,

$$\mathbf{p}^* = (\mathbf{s}^*)^T \hat{\boldsymbol{\beta}}_k. \quad (3)$$

The clear advantage of this process, when successful, is that it removes the need to perform laboratory measurements. The price to be paid is the need to solve the (highly) underdetermined system of equations (1) with $n \gg m$.

A variety of methods, including principal component analysis [7], partial least squares (PLS) [13] and independent component analysis [12], have been utilised to perform the dimension reduction that is the essence of the calibration. Independently, because of the high accuracy with which NIR spectra are recovered using computer controlled instrumentation, it has been established that derivative spectroscopy techniques (section 2) can be utilised to explore differences in molecular structure of barley mutants [16]. This leads naturally to

1. whether, for the calibration step, there might be a more appropriate choice than the measured NIR spectra, such as the fourth derivative of the NIR spectra, and, hence,
2. the discussion in section 3 of the appropriateness of preprocessing the measured spectral data before performing the calibration step.

Whatever the form of the underdetermined system to be solved, some appropriate dimension reduction methodology is usually applied. A popular strategy is PLS [13]. It exploits, using an iterative covariance maximisation strategy, the fact that any matrix can be rewritten as the sum of rank-one matrices. The essence of the PLS strategy reduces to finding iteratively the matrix of ‘factors’ $F \in \mathbb{R}^{m \times k}$, with $k < n$, which generates the following rank-one representation for S and the corresponding decomposition of the property vector \mathbf{p} :

$$S = FQ^T + E, \quad \mathbf{p} = F\mathbf{c} + \mathbf{r}, \quad (4)$$

with the residuals E and \mathbf{r} appropriately small.

As noted by Anderssen et al. [6], the PLS strategy represents an implementation for the following ‘*simultaneous minimisation*’ procedure: choose the factors F to minimise, for some appropriate choice of norm,

$$\lambda \|FQ^T - S\|^2 + (1 - \lambda) \|F\mathbf{c} - \mathbf{p}\|^2, \quad (5)$$

with λ controlling the trade-off between the fits to S and \mathbf{p} . The germ of this idea was proposed and analysed by Frank as *biased regression* [8] and *intermediate least squares regression* [9]. The consequences of this interpretation are pursued in section 4, where it is shown that this *simultaneous minimisation* is equivalent to solving the total least squares problem, in the Frobenius norm, for a matrix augmentation of $S\boldsymbol{\beta} = \mathbf{p}$.

2 Introductory background

In order to motivate the matters highlighted in the Introduction and the deliberations to follow in sections 3 and 4, it is necessary to give some appropriate background.

2.1 NIR spectroscopy

The measurement of the spectroscopic response of a material to an electromagnetic stimulus, such as the vibrational response of a biomaterial to a near-infrared (NIR) stimulus, records indirectly the proportional presence of the molecular components in the material that respond to the stimulus. For biomaterials such as cereals, the NIR spectrum records the intensity of vibration of the size chains on the protein molecules in the cereals. Because the wavelengths at which the various side chains vibrate are known, the wavelengths, and hence the sidechains, become the molecular identifiers for the proteins present and their proportional presences. This information recovery protocol has been used to examine successfully the proportional presence of various proteins such as gliadin and glutenin in wheat [15] and the molecular difference in barley mutants [16]. An introduction to the background science of NIR spectroscopy and its application can be found in the book of Osborne et al. [14].

2.2 Calibration and prediction methodologies

PLS is one of a large number of procedures available for solving underdetermined systems of equations. Elden [7] gave an insightful discussion about the relationship between PLS and singular value decomposition (for which PCR corresponds to an implementation).

As explained by Gosselin et al. [11], the calibration should be *“a simple wavelength selection method, yet having the ability to identify relevant spectral intervals.”* In relation to our SVD analysis of wheat NIR spectra, it is with this interval aspect of calibration that preprocessing-with-differentiation has a potential role to play.

2.3 Derivative spectroscopy

It is well known that the (numerical) differentiation of observational data enhances the presence of the measurement noise [1, 3]. Equally important, known historically since before the 1920s and exploited in analytic chemistry, differentiation also has the potential to enhance the fine scale structure in observational data [5, 16]. The accuracy and detail with which the data have been measured are the factors determining whether the enhancement of the measurement errors or the fine scale will dominate when the differentiation is performed.

It is only with the advent of modern computer controlled instrumentation, such as NIR instruments, that the full potential of derivative spectroscopy has materialised, since it allows, through a multiple sweeping and averaging process, highly accurate data on very fine wavelength grids to be generated. Anderssen and Hegland [5] analysed in detail the trade-offs involved.

3 Preprocessing of the spectra

As explained above, the goal of the calibration step is the estimation of $\boldsymbol{\beta}$ as $\hat{\boldsymbol{\beta}}_k$, corresponding to a dimension reduction. Those components of $\hat{\boldsymbol{\beta}}_k$ will be essentially zero which multiply the spectral responses at wavelengths that are not excited by the molecular vibrations associated with the property \mathcal{P} .

However, the choice of $\boldsymbol{\beta}$ in the calibration equation (1) to define how the prediction process (of equation (3)) should be performed is not unique. This is directly reflected in that the calibration equation (1) can be rewritten to take the alternative form

$$SGG^{-1}\boldsymbol{\beta} = \mathbf{p}, \quad \bar{S}\bar{\boldsymbol{\beta}} = \mathbf{p}, \quad \bar{S} = SG, \quad \bar{\boldsymbol{\beta}} = G^{-1}\boldsymbol{\beta}, \quad (6)$$

where the matrix G represents the chosen preprocessing to be performed on each of the spectra forming S . In NIR spectroscopy, in order to remove the

linear scattering sample effect, second order differentiation preprocessing is often applied [14]. If one-sided derivatives are included in the differentiation of \mathbf{S} , then \mathbf{G} would correspond to the second derivative matrix \mathbf{D}^2 . An alternative choice for \mathbf{G} is the fourth differentiation matrix \mathbf{D}^4 [16]. It is also justified because the vibrational spectroscopic response to an NIR stimulus is strongly localised and not global polynomial in nature.

In this preprocessing situation, the modified calibration equations (6) have the same basic structure as the original calibration equations (1). Consequently, in the calibration step, the estimation of $\boldsymbol{\beta}$ as $\hat{\boldsymbol{\beta}}_k$ is replaced by the estimation of $\bar{\boldsymbol{\beta}}$ as $\hat{\bar{\boldsymbol{\beta}}}_k$. In this way, the prediction of the value of \mathcal{P} as \mathbf{p}^* using the predictor (3) is replaced by the predictor

$$\bar{\mathbf{p}}^* = (\mathbf{s}^* \mathbf{G})^T \hat{\bar{\boldsymbol{\beta}}}_k. \quad (7)$$

3.1 The fourth differentiation possibility

In order to motivate the use of the fourth derivative, the following matters will be examined:

1. a theoretical analysis;
2. an SVD analysis of wheat NIR data.

3.1.1 Theoretical considerations

An SVD analysis of the matrices \mathbf{S} and $\bar{\mathbf{S}}$ yields

$$\mathbf{S} = \mathbf{U} \boldsymbol{\Sigma} \mathbf{V}^T = \sum_{j=1}^m \mathbf{u}_j \sigma_j \mathbf{v}_j^T, \quad \bar{\mathbf{S}} = \bar{\mathbf{U}} \bar{\boldsymbol{\Sigma}} \bar{\mathbf{V}}^T = \sum_{j=1}^m \bar{\mathbf{u}}_j \bar{\sigma}_j \bar{\mathbf{v}}_j^T, \quad (8)$$

$$\text{and } \boldsymbol{\beta} = \sum_{j=1}^m \frac{\mathbf{u}_j^T \mathbf{p}}{\sigma_j} \mathbf{v}_j, \quad \hat{\boldsymbol{\beta}} = \sum_{j=1}^m \frac{\bar{\mathbf{u}}_j^T \mathbf{p}}{\bar{\sigma}_j} \bar{\mathbf{v}}_j, \quad (9)$$

where the $(\mathbf{m} \times \mathbf{m})$ -matrices \mathbf{U} and $\bar{\mathbf{U}}$ are orthogonal, the $(\mathbf{m} \times \mathbf{n})$ -matrices \mathbf{V}^T and $\bar{\mathbf{V}}^T$ are the orthogonal columns of \mathbf{V} and $\bar{\mathbf{V}}$ and the $(\mathbf{m} \times \mathbf{m})$ -matrices Σ and $\bar{\Sigma}$ are diagonal matrices of the (positive) singular values with their entries arranged in decreasing order. Formally, as shown in equation (8), the SVD decomposition consists of a sum of rank-one components. As indicated in equation (4), PLS is also constructed as a sum of rank-one matrices. They are generated iteratively in PLS taking the structure in \mathbf{p} into account as an essential part of the iteration.

In an SVD analysis, a comparison of the rate of decay of the singular values is known to be a useful characterisation of the degree of linear independence between the rows of an underdetermined matrix [10]. It therefore represents a basis for the comparison of \mathbf{S} and $\bar{\mathbf{S}}$. Equally important information is obtained by comparing the structure of the first, second, etc., rank-one components.

3.1.2 SVD analysis of wheat NIR data

Figure 1 plots the wheat NIR spectra for 59 representative wheat samples [2] and their fourth derivatives.

Their first rank-one components are not plotted as they are quite similar to those shown in Figure 1. Their second and third rank-one components are plotted in Figures 2 and 3, respectively.

The advantage of using the fourth derivative has become clear. The second and third rank-one corrections for both the NIR spectra and their fourth derivatives contain localised features which identify where the NIR spectra see differences in the response of the samples to the NIR stimulus. The identification is much stronger and clearer in the fourth derivatives than the original spectra. It yields an indicative proof that, under appropriate conditions, the differentiation yields information about the fine scale structure which often has a strong linkage to the property being calibrated.

As observed by others, and explained succinctly by Elden [7], the success of

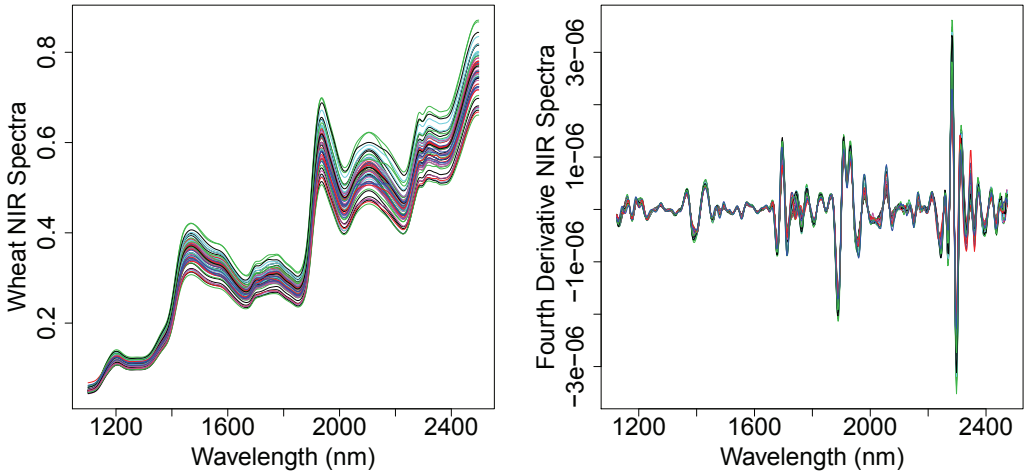


FIGURE 1: The NIR spectra of the 59 wheat samples (left) and their fourth derivatives (right).

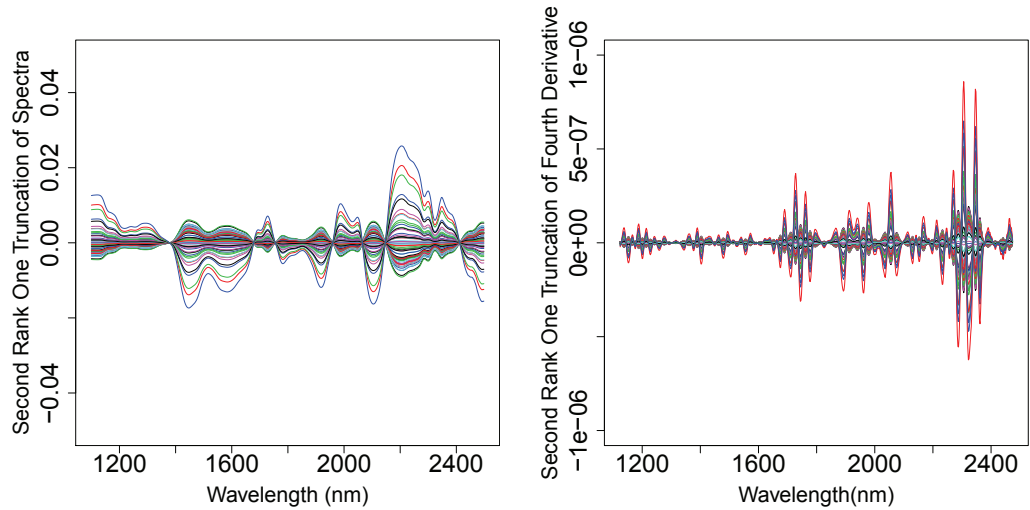


FIGURE 2: The second rank-one SVD component of the NIR spectra of the 59 wheat samples (left) and their fourth derivatives (right).

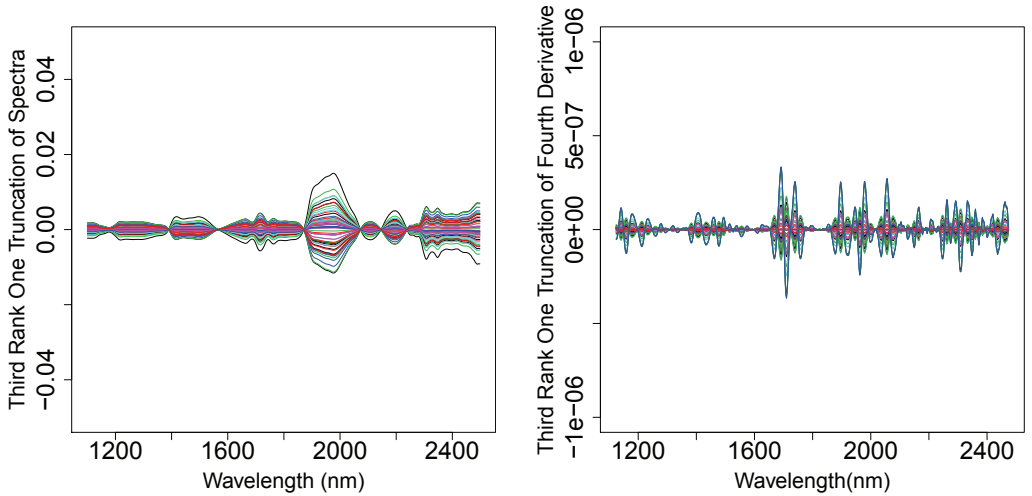


FIGURE 3: The third rank-one SVD component of the NIR spectra of the 59 wheat samples (left) and their fourth derivatives (right).

PLS compared with PCR is the iterative process of the former which takes account of the structure in the right hand side \mathbf{p} of the calibration equation (1) with respect to the successive choices of the components used to form the dimension reduction. This process is implicitly encapsulated in the structure of the simultaneous minimisation representation of equation (5). It thereby represents an alternative theoretical framework in which to explore rank-one dimension reduction.

Elden [7] suggests that, instead of simply choosing the rank-one corrections on the basis of the sizes of their singular values, PCR should be modified to choosing the rank-one corrections on the basis of the sizes of $\mathbf{u}_j^T \mathbf{p}$, the SVD projections onto \mathbf{p} . For the NIR data, Table 1 gives the first seven projections for the original, second derivative and fourth derivative data.

As is clear from the values of the entries in Table 1, the first two SVD rank-one components for the original data and the first three for the fourth derivative data account for about 99% of the variability. This difference in the number

TABLE 1: SVD projection protein functionals

	NIR spectra	Second derivative spectra	Fourth derivative spectra
1	-95.3267072	-94.6710192	-94.03540423
2	33.60960857	34.54571042	30.57111018
3	-3.838051909	-8.878968064	-21.37190323
4	3.172797189	-2.435736995	0.325186499
5	-1.076245832	-1.398356549	-1.621886793
6	3.125263444	-0.263572285	-1.760102019
7	2.211280123	-0.695537032	2.212543474

of rank-one components required to explain the bulk of the structure in the original and fourth derivative data represents algebraic confirmation of the structure seen in Figures 2 and 3. In addition, it illustrates, in terms of the required third rank-one component, how the differentiation has identified fine scale structure in the original data which its SVD analysis has failed to highlight.

This shows that there is key localised structure in the NIR wheat spectra which the fourth derivative identifies more strongly than either the original data or the second derivative. This thereby gives a good illustration of the utility of performing the calibration with the fourth derivative.

Independent suggestive confirmation of the localisation performed by the differentiation is shown in Figure 4 where slightly smoothed values of β_k and $\bar{\beta}_k$ of equation (9) are plotted for $m = 3$.

4 Simultaneous minimisation

As mentioned above, the importance of the simultaneous minimisation of equation (5) is that it represents an alternative framework for exploring rank-one dimension reduction. As already mentioned, the germ of this idea

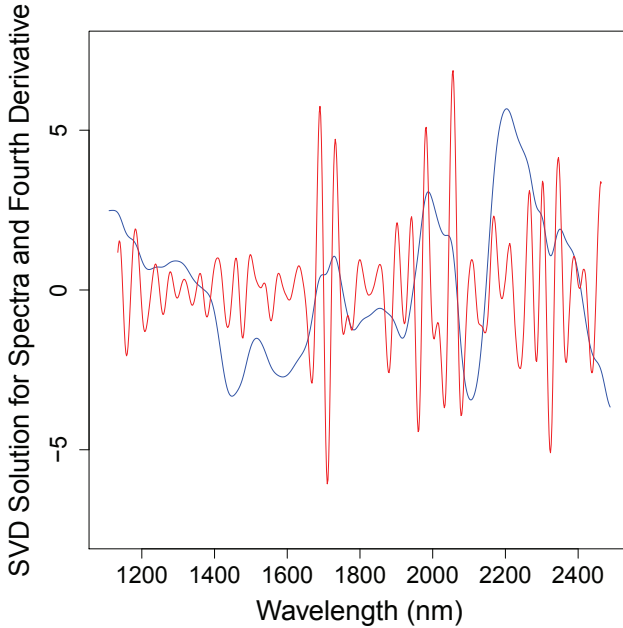


FIGURE 4: Slightly smoothed approximations for β (blue) and $\hat{\beta}$ (red) for the NIR wheat data, using the sum of the first three rank-one components.

goes back to Frank [8, 9].

Here, we prove the following proposition.

Proposition 1 *The simultaneous minimisation of equation (5) is equivalent to solving the total least squares problem $\min_{\mathcal{F}} \|\hat{\mathbf{S}} - \mathbf{F}\hat{\mathbf{Q}}^T\|_{\mathcal{F}}^2$, where $\hat{\mathbf{S}}$ and $\hat{\mathbf{Q}}$ denote respectively the augmented matrices $[\mathbf{S}, \theta\mathbf{p}]$ and $[\mathbf{Q}, \theta\mathbf{c}]$.*

Proof: Using the trace property of the Frobenius norm, one obtains, using the linearity of the trace operator,

$$\begin{aligned} \|\hat{\mathbf{S}} - \mathbf{F}\hat{\mathbf{Q}}^T\|_{\mathcal{F}}^2 &= \text{trace}((\hat{\mathbf{S}} - \mathbf{F}\hat{\mathbf{Q}}^T)^T(\hat{\mathbf{S}} - \mathbf{F}\hat{\mathbf{Q}}^T)) \\ &= \text{trace}(\hat{\mathbf{S}}^T\hat{\mathbf{S}}) - \text{trace}((\hat{\mathbf{S}})^T(\mathbf{F}\hat{\mathbf{Q}}^T)) \end{aligned}$$

$$- \text{trace}((\mathbf{F}\hat{\mathbf{Q}}^T)(\hat{\mathbf{S}})) + \text{trace}((\mathbf{F}\hat{\mathbf{Q}}^T)^T(\mathbf{F}\hat{\mathbf{Q}}^T)).$$

On using the following identities

$$\begin{aligned} \text{trace}((\hat{\mathbf{S}})^T(\hat{\mathbf{S}})) &= \text{trace}(\mathbf{S}^T\mathbf{S}) + \theta^2\mathbf{p}^T\mathbf{p}, \\ \text{trace}((\hat{\mathbf{S}})^T(\mathbf{F}\hat{\mathbf{Q}}^T)) &= \text{trace}(\mathbf{S}^T(\mathbf{F}\mathbf{Q}^T)) + \theta^2\mathbf{p}^T(\mathbf{F}\mathbf{c}), \\ \text{trace}((\mathbf{F}\hat{\mathbf{Q}}^T)(\hat{\mathbf{S}})) &= \text{trace}((\mathbf{F}\mathbf{Q}^T)^T\mathbf{S}) + \theta^2(\mathbf{F}\mathbf{c})^T\mathbf{p}, \\ \text{and } \text{trace}((\mathbf{F}\hat{\mathbf{Q}}^T)^T(\mathbf{F}\hat{\mathbf{Q}}^T)) &= \text{trace}((\mathbf{F}\mathbf{Q}^T)^T(\mathbf{F}\mathbf{Q}^T)) + \theta^2(\mathbf{F}\mathbf{c})^T(\mathbf{F}\mathbf{c}), \end{aligned}$$

it follows that

$$\begin{aligned} \|\hat{\mathbf{S}} - \mathbf{F}\hat{\mathbf{Q}}^T\|_{\mathcal{F}}^2 &= \|\mathbf{S} - \mathbf{F}\mathbf{Q}^T\|_{\mathcal{F}}^2 + \theta^2(\mathbf{p}^T\mathbf{p} - \mathbf{p}^T(\mathbf{F}\mathbf{c}) - (\mathbf{F}\mathbf{c})^T\mathbf{p} + (\mathbf{F}\mathbf{c})^T(\mathbf{F}\mathbf{c})) \\ &= \|\mathbf{S} - \mathbf{F}\mathbf{Q}^T\|_{\mathcal{F}}^2 + \theta^2\|\mathbf{p} - \mathbf{F}\mathbf{c}\|_2^2. \end{aligned}$$



With $\theta^2 = (1 - \lambda)/\lambda$, the right hand side of the last equation in the proof has a structure indicative of the simultaneous minimisation formulation of equation (5).

References

- [1] B. Anderssen, F. de Hoog, and M. Hegland. A stable finite difference ansatz for higher order differentiation of non-exact data. *Bull. Austral. Math. Soc.*, 58(2):223–232, 1998. [C339](#)
- [2] R. S. Anderssen, E. Carter, B. G. Osborne, and I. J. Wesley. Joint inversion of multi-modal spectroscopic data of wheat flours. *Appl. Spectro.*, 59(7):920–925, JUL 2005. [C335](#), [C341](#)
- [3] R. S. Anderssen and F. R. de Hoog. Finite-difference methods for the numerical differentiation of non-exact data. *Computing*, 33(3–4):259–267, 1984. [C339](#)

- [4] R. S. Anderssen and R. Haraszi. Characterizing and exploiting the rheology of wheat hardness. *Euro. Food Res. Tech.*, 229(1):159–174, MAY 2009. doi:10.1007/s00217-009-1037-9 C335
- [5] R. S. Anderssen and M. Hegland. Derivative Spectroscopy—An enhanced role for numerical differentiation. *J. Integ. Eqn. Appl.*, 22(3):355–367, 2010. doi:10.1216/JIE-2010-22-3-355 C339
- [6] R. S. Anderssen, B. G. Osborne, and I. J. Wesley. The application of localisation to near infrared calibration and prediction through partial least squares regression. *JNIRS*, 11(1):39–48, 2003. C337
- [7] L. Elden. Partial least-squares vs. Lanczos bidiagonalization—I: analysis of a projection method for multiple regression. *Comp. Stats & Data Anal.*, 46(1):11–31, 2004. doi:10.1016/S0167-9473(03)00138-5 C336, C338, C341, C343
- [8] I. E. Frank. Beyond linear least-squares regression. *Trac-Trends in Anal. Chem.*, 6(10):271–275, 1987. C337, C345
- [9] I. E. Frank. Intermediate least-squares regression method. *Chemo. Intel. Lab. Systems*, 1(3):233–242, 1987. C337, C345
- [10] G. H. Golub and C. Reinsch. Singular value decomposition and least squares solutions. *Numer. Math.*, 14(5):403–420, 1970. C341
- [11] R. Gosselin, D. Rodrigue, and C. Duchesne. A Bootstrap-VIP approach for selecting wavelength intervals in spectral imaging applications. *Chemo. Intell. Lab. Systems*, 100(1):12–21, 2010. doi:10.1016/j.chemolab.2009.09.005 C338
- [12] A. Hyvarinen and E. Oja. Independent component analysis: algorithms and applications. *Neural Net.*, 13(4–5):411–430, 2000. C336
- [13] T. Naes, T. Isaksson, T. Fearn, and T. Davies. *A User-Friendly Guide to Multivariate Calibration and Classification*. NIR Publications, Chichester, UK, 2002. C336, C337

- [14] B. G. Osborne, T. Fearn, and P. H. Hindle. *Practical NIR Spectroscopy with Applications in Food and Beverage Analysis*. Longman Scientific and Technical, Harlow, UK, 1993. McGraw-Hill Series in Higher Mathematics. [C335](#), [C336](#), [C338](#), [C340](#)
- [15] I. J. Wesley, O. Larroque, B. G. Osborne, N. Azudin, H. Allen, and J. H. Skerritt. Measurement of gliadin and glutenin content of flour by NIR spectroscopy. *J. Cereal Sci.*, 34(2):125–133, 2001. [C338](#)
- [16] P. R. Wiley, G. J. Tanner, P. M. Chandler, and R. S. Anderssen. Molecular classification of barley mutants using derivative spectroscopic analysis of NIR spectra of their wholemeal flours. *J. Agri. Food Chem*, (2), 2009. [C336](#), [C338](#), [C339](#), [C340](#)

Author addresses

1. **R. S. Anderssen**, CSIRO Mathematics, Informatics and Statistics
GPO Box 664, Canberra, ACT 2601, AUSTRALIA.
<mailto:Bob.Anderssen@csiro.au>
2. **F. R. de Hoog**, CSIRO Mathematics, Informatics and Statistics
GPO Box 664, Canberra, ACT 2601, AUSTRALIA.
<mailto:Frank.deHoog@csiro.au>
3. **I. J. Wesley**, BRI Australia, 1 Rivett Road, North Ryde, NSW 2113,
AUSTRALIA.
<mailto:I.Wesley@bri.com.au>