# PLS pour les données compositionnel appliqué à la texture du sol et la composition du raisin

# Partial Least Squares for compositional data applied to soil texture and grape composition



## 1. Introduction

Soil has been acknowledged as one of the main environmental factors governing grape and resulting wine characteristics (Van Leeuwen and Seguin, 2006). For that, along with climate, it is a major element of the France concept of *terroir* (Morlat and Bodin, 2006). Soil properties may affect grape characteristics to a different extend, depending on the pedo-environmental specificity a vineyard district. Thus, the knowledge of the relationships between soil properties and grape characteristics becomes a key element for sustainable vineyard management in that district.

Among soil properties, texture plays a major role (White, 2003). Texture refers to the proportions of different-sized particles in a soil. To study the mineral particles of a soil, scientists separate them into groups according to size. There are different scheme for classifying soil particles. In all classification schemes, the sum of all particles (by weight) less than or equal to 2 mm is equal to 100% and is referred as the "fine-earth fraction". The texture of a soil is not readily subject to change, so it is considered a basic property of a soil (Brady and Weil, 2002).

Texture is soil's most important physical property for grapegrowing, because of its influence on other soil properties, crucial to plant growth, such as water intake rates (infiltration), water movements through soil (hydraulic conductivity), soil water holding capacity, the easy of tilling the soils, and the amount of aeration (which is vital to root growth). Texture also influence soil fertility, and overall vine vigor (Rice et al., 2002).

Based on the above consideration, it would be important to investigate the relationships between soil texture and grape in term of relative magnitudes within the boundaries of a vineyard district, where the grape composition (**Y**) can be described as a function of the soil texture (**X**). Frequently extract the information in the response matrix **Y** by means of descriptor matrix **X** is done using Partial Least Squares regression (Wold et al., 2001) as alternative to ordinary least-squares regression. Both explicative and response matrix present compositional data where these particular properties give origin to special problems for imputation, and they can rarely be analyzed with traditional PLS. In fact the original data values require transforming in order to depict correctly the structures that are appropriate to the particular nature of the compositional data. In mathematical terms, a vector $\tilde{\mathbf{w}} = (\tilde{w}_1, \ldots, \tilde{w}_J)$ is compositional if the elements $\tilde{w}_j \geq 0$ (for $j = 1, \ldots, J$) and $\tilde{w}_1 + \ldots + \tilde{w}_J = 1$.

Compositional data are also special in this respect and careful consideration of the relationships between parts of a composition is required before we embark on applying PLS analysis. In literature, Hinkle and Rayner (1995) suggested to use a PLS for study the relationship between a matrix of compositional data and one response variable. Following the Aitchson's approach (1986) to investigate the relationships between soil texture and grape composition in the Telesina Valley vineyard district (Southern Italy) a double logcontrast transformation on is proposed.

After a preliminarily discussion on the compositional data, the PLS for **X** and **Y** compositional data is given. Finally, principal results of the relationship between soil texture and grape are presented.

# 2. Theory

## 2.1 Compositional data

Define $\tilde{w}_1,\ldots,\tilde{w}_J$ as positive quantities with the same measurement scale $\tilde{\mathbf{w}} = (\tilde{w}_1,\ldots,\tilde{w}_J)$ $\tilde{w}_1 \geq 0,\ldots,\tilde{w}_J \geq 0$ and $\|\tilde{\mathbf{w}}\|$ the trace of $\tilde{\mathbf{w}}$. The vector $\tilde{\mathbf{w}}$ is the basis of compositional data and $\mathbf{w} = \tilde{\mathbf{w}}/\|\tilde{\mathbf{w}}\|$ is a composition vector. Moreover two bases $\tilde{\mathbf{w}}$ and $\dot{\mathbf{w}}$ are compositional equivalents if there exists a positive constant *h* such that $\tilde{\mathbf{w}} = h\,\dot{\mathbf{w}}$. This equivalence relation partitions the space in equivalence classes, called compositions.

More generally, we define $\mathbf{W}$ $(N, J)$ a compositional data matrix if all elements are positive and each row is constrained to the unit-sum $\mathbf{W}\hat{\mathbf{1}}_J = \hat{\mathbf{1}}_N$ where $\hat{\mathbf{1}}_J$ and $\hat{\mathbf{1}}_N$ are vectors of units of *J* and *N* dimension, respectively. Let $\mathbf{Q} = \left[N\mathbf{I}_N - \hat{\mathbf{1}}_N\hat{\mathbf{1}}_N^t\right]$ be the product between $1/N$ and the usual centering projector then $\mathbf{W}^t\mathbf{QW}$ is the covariance matrix of $\mathbf{W}$ called *crude covariance matrix* (Aitchison, 1986). The unit-sum constraint for each row of $\mathbf{W}$ implies four difficulties: 1) Negative bias, 2) Subcomposition, 3) Basis, 4) Null correlation. Each row and column of $\mathbf{W}^t\mathbf{QW}$ has zero-sum: $\mathbf{W}^t\mathbf{QW}\hat{\mathbf{1}}_J = \hat{\mathbf{0}}_J$ where $\hat{\mathbf{0}}_J$ is a *J* dimensional vector of zero. Therefore each variable has a covariance sum equivalent to negative variance (the first difficulty). No-relationship exists between the crude covariance matrix and the crude subcomposition covariance one. Therefore the variation of subcomposition can substantially influence the covariance (the second difficulty). Likewise in the subcomposition, it is not easy to select a basis $\tilde{\mathbf{w}}$ for the composition (which is the third difficulty). Like the *crude covariance matrix*, each row and column of the *crude correlation matrix* of $\mathbf{W}$ has a zero-sum. Therefore the correlation between two variables is not free to range over the usual interval $[-1,1]$. The negative bias causes a radical difference from the standard interpretation of correlation between variables. Zero correlation between two ratios does not mean that there is no association (the latter difficulty). Moreover the uninterpretable crude covariance structure is not the only problem of compositional data. Unfortunately, compositional data often exhibit curvature when standard multivariate methods are employed.

Aitchison (1986) richly described the properties of compositional data and proposed an alternative form of logratio, where the more useful is based a geometric mean $g(\mathbf{w})$. Replacing the natural non-negative condition by the following stronger assumption of the strict positive quantities: $\tilde{w}_{i1} > 0, \dots, \tilde{w}_{iJ} > 0$. This assumption in economics or marketing science is a problem (Gallo, 2003). Nevertheless, as it is fully compatible with the nature of soil texture and grape composition data sets. Assuming the strict positive quantities, Aitchison (1982) proposes to transform each element of $\mathbf{W}$ ($w_{ij}$) in the logratio $\log\left[w_{ij}/g(\mathbf{w})\right]$, that because the relative matrix of centred logratio $\mathbf{Z}$, with generic element $z_{ij} = \log\left[w_{ij}/g(\mathbf{w})\right]$ is adequate for a low-dimensional description of compositional variability. Moreover, a generalization of the logratios – called *logcontrasts* – have particular and researched proprieties in compositional data analysis. Logcontrast of $\mathbf{w}$ is any loglinear combination $\mathbf{c}' \log \mathbf{w} = c_1 \log w_1 + \dots + c_J \log w_J$ with $c_1 + \dots + c_J = 0$, where of logcontrast with the geometric mean $g(\mathbf{w})$ presents the property: $\mathbf{c}' \log \mathbf{w} = \mathbf{c}' \log(\mathbf{w}/g(\mathbf{w}))$.

## 2.2 PLS for compositional data

Let $\mathbf{A} = \left[\mathbf{a}_1, \dots, \mathbf{a}_h\right]$ and $\mathbf{B} = \left[\mathbf{b}_1, \dots, \mathbf{b}_h\right]$ the first *h* coefficient vectors with $\mathbf{a} \in \Re^J$ and $\mathbf{b} \in \Re^K$, the aim of PLS is to search components $\mathbf{t}_{h+1} = \mathbf{X}\mathbf{a}_{h+1}$ and $\mathbf{u}_{h+1} = \mathbf{Y}\mathbf{b}_{h+1}$ that maximize

$\operatorname{cov}^2\left(\mathbf{Xa}_{h+1}, \mathbf{Yb}_{h+1}\right)$ with normalizations and orthogonal constraints respectively. Maximize $\operatorname{cov}^2\left(\mathbf{Xa}_{h+1}, \mathbf{Yb}_{h+1}\right)$ with constrained $\|\mathbf{a}\|^2 = 1$, $\|\mathbf{b}\|^2 = 1$ and $\mathbf{t}_{h+1}^t \mathbf{T} = \hat{\mathbf{0}}$ with $\mathbf{T} = \left[\mathbf{t}_1, \ldots, \mathbf{t}_h\right]$ is a eigenstructure problem that can be solved using Lagrange multipliers. The Lagrangian approach becomes particularly difficult and unintuitive when PLS scores are constrained to be uncorrelated. Raynes (2000) shows how any constraint of this type can be viewed as an orthogonality constraint, with respect to a redefined inner product. And relative inner supremum is a simple principal components problem.

Suppose $\mathbf{A} = \left[\mathbf{a}_1, \ldots, \mathbf{a}_h\right]$ and $\mathbf{B} = \left[\mathbf{b}_1, \ldots, \mathbf{b}_h\right]$ the first $h$ coefficient vectors and $\mathbf{Q} = \left[N\mathbf{I}_N - \hat{\mathbf{1}}_N \hat{\mathbf{1}}_N^t\right]$ is the product between $1/N$ and the usual centering projector then $\mathbf{\Sigma}_X = \mathbf{X}'\mathbf{QX}$ is the covariance matrix of $\mathbf{X}$, $\mathbf{\Sigma}_Y = \mathbf{Y}'\mathbf{QY}$ is the covariance matrix of $\mathbf{Y}$

$$\underset{\substack{\mathbf{a}\in\Re^J;\ \mathbf{b}\in\Re^K \\ \mathbf{a}^t\Sigma_X\mathbf{A}=\hat{\mathbf{0}} \\ \mathbf{b}^t\Sigma_Y\mathbf{B}=\hat{\mathbf{0}}}}{\arg\max} \left\{\frac{\operatorname{cov}^2(\mathbf{Xa}, \mathbf{Yb})}{\left(\mathbf{a}^t\mathbf{a}\right)\left(\mathbf{b}^t\mathbf{b}\right)}\right\} = \left\{\mathbf{a}_{h+1}, \mathbf{b}_{h+1}\right\} \tag{1}$$

$\mathbf{a}_{h+1}$ corresponding to the largest eigenvalue $\lambda_{h+1}$ of $\left(\mathbf{I}_J - \mathbf{P}_h^X\right)\Sigma_{XY}\left(\mathbf{I}_K - \mathbf{P}_h^Y\right)\Sigma_{YX}$ where

$\Sigma_{XY} = \mathbf{X}^t\mathbf{QY}$, $\Sigma_{YX} = \mathbf{Y}^t\mathbf{QX}$, $\mathbf{P}_h^X = \left(\Sigma_X\mathbf{A}\left(\mathbf{A}^t\Sigma_X\Sigma_X\mathbf{A}\right)^{-1}\mathbf{A}^t\Sigma_X\right)$, $\mathbf{P}_h^Y = \left(\Sigma_Y\mathbf{B}\left(\mathbf{B}^t\Sigma_Y\Sigma_Y\mathbf{B}\right)^{-1}\mathbf{B}^t\Sigma_Y\right)$,

and $\mathbf{b}_{h+1} = \left(\mathbf{I}_J - \mathbf{P}_h^X\right)\Sigma_{YX}\mathbf{a}_{h+1}$.

Let $\mathbf{Z}$ $(\mathrm{N}, \mathrm{J})$ and $\mathbf{V}$ $(\mathrm{N}, G)$ be the matrix of a centred logratio with one-sum for each row and column. The covariance matrix $\Sigma_Z$ has the property $\Sigma_Z\hat{\mathbf{1}}_J = \hat{\mathbf{0}}_\mathrm{J}$ and $\mathbf{P}_J^\perp\Sigma_Z = \Sigma_Z$ where $\mathbf{P}_J^\perp = \left(\mathbf{I}_J - \hat{\mathbf{1}}_J^t\hat{\mathbf{1}}_J/J\right)$.

Similarly, the covariance matrix $\Sigma_V$ has the property $\Sigma_V\hat{\mathbf{1}}_K = \hat{\mathbf{0}}_K$ and $\mathbf{P}_K^\perp\Sigma_V = \Sigma_V$ where

$\mathbf{P}_K^{\perp} = \left(\mathbf{I}_K - \hat{\mathbf{1}}_K^t \hat{\mathbf{1}}_K / K\right)$. Every linear combination of $\mathbf{z}$ and $\mathbf{v}$ is equivalent to zero, because $\mathbf{Z}\mathbf{P}_J^{\perp}\mathbf{a} = \mathbf{Z}\mathbf{a}$ and $\mathbf{V}\mathbf{P}_K^{\perp}\mathbf{b} = \mathbf{V}\mathbf{b}$ then the vectors $\mathbf{a} = \left[\mathrm{a}_1, \ldots, \mathrm{a}_J\right]$ and $\mathbf{b} = \left[b_1, \ldots, b_K\right]$ have zero-sum: $\mathbf{a}^t \hat{\mathbf{1}}_J = 0$ and $\mathbf{b}^t \hat{\mathbf{1}}_K = 0$. Aim of PLS on compositional data is to maximize $\operatorname{cov}^2\left(\mathbf{Za}, \mathbf{Yb}\right)$ subject to the usual constraints of PLS plus the additional constraints that each coefficient vector has zero-sum: $\mathbf{a}^t \hat{\mathbf{1}}_J = 0$ and $\mathbf{b}^t \hat{\mathbf{1}}_K = 0$.

To obtain the solutions we define $\mathbf{A} = \left[\mathbf{a}_1, \ldots, \mathbf{a}_h\right]$ and $\mathbf{B} = \left[\mathbf{b}_1, \ldots, \mathbf{b}_h\right]$ the first $h$ coefficient vectors (see appendix),

$$\underset{\substack{\mathbf{a}\in\Re^J;\ \mathbf{b}\in\Re^K \\ \mathbf{a}^t\Sigma_Z\mathbf{A}=\hat{\mathbf{0}} \\ \mathbf{b}^t\Sigma_V\mathbf{B}=\hat{\mathbf{0}} \\ \mathbf{a}^t\hat{\mathbf{1}}_J=0;\ \mathbf{b}^t\hat{\mathbf{1}}_K=0}}{\arg\max} \left\{\frac{\operatorname{cov}^2\left(\mathbf{Za}, \mathbf{Vb}\right)}{\left(\mathbf{a}^t\mathbf{a}\right)\left(\mathbf{b}^t\mathbf{b}\right)}\right\} = \underset{\substack{\mathbf{a}\in\Re^J;\ \mathbf{b}\in\Re^K \\ \mathbf{a}^t\Sigma_Z\mathbf{A}=\hat{\mathbf{0}} \\ \mathbf{b}^t\Sigma_V\mathbf{B}=\hat{\mathbf{0}}}}{\arg\max} \left\{\frac{\operatorname{cov}^2\left(\mathbf{Za}, \mathbf{Vb}\right)}{\left(\mathbf{a}^t\mathbf{a}\right)\left(\mathbf{b}^t\mathbf{b}\right)}\right\} = \left\{\mathbf{a}_{k+1}, \mathbf{b}_{k+1}\right\} \qquad (2)$$

# 3. Application

## 3.1. The study area

The Telesina Valley study area is located in the western Benevento province, Campania region, Southern Italy, within the boundaries of the middle lower basin of the Calore river, upstreams of the confluence of the latter with the Volturno river. The morpho-structural setting of the area is defined by the graben of the Calore river Valley delimited to the north and south by the calcareous horst of Matese-Monte Maggiore and the Taburno-Camposauro mountain groups, respectively, both of them extending from east to west of the investigated area. Agricultural land use is dominated by vineyards which cover about 6.500 ha (i.e., 25% of the study area).

Viticulture is practised on eight different landscape units, namely footslope on alluvial-colluvial and scree deposits of Taburno-Camposauro mountain range, floodplain, alluvial terraces, terraces of complex genesis, intermediate slope belt of alluvial-colluvial aggradation, pediment, sandstone and marly hills of Titerno river, terraces on the Campanian ignimbrite. The main soil types associated with the above landscape are: *typic* and *fluventic Haplustepts*, *typic Calciustepts*, *typic* and *humic Ustifluvents*, *typic* and *vitrandic Haplustolls*, *typic* and *vitrandic Calciustolls*, *alfic Ustivitrands*, *vitrandic Haplustals*.

## 3.2 Soil and grape sampling and analysis

The study was conducted on the Falangina cultivar, one the most celebrated white grapes cultivar of study area and, more generally, of the Campania region (Monaco *et al.*, 2004). Eight nine sites were selected in such a manner to represent, as much as possible, the variability of the vineyard landscape. Sampling locations were recorded using a Garmin 12 channels GPS and adjusted through visual identification of the sites on 1:10.000 georeferenced colour orthophotographs, under Arc-View GIS 3.2 environment. The selection of sites was conditioned by both the effective spatial distribution of Falanghina vineyards.

At each sampling site, ten representative vines falling in a unique rectangular plot were selected. At harvest, a representative number of grape clusters was collected following the methodology proposed by Moio et al. (1999). Berries were pressed for the extraction of juice which was analysed for titratable acidity (TAc), malic acid (Mal), tartaric acid (Tar), total soluble solids (Brix ), and pH (GpH). TAc, Brix, and GpH were determined according to the Official Analytical Methods (MAF, 1985); Mal was determined using the enzymatic method (Boehringer, 1983); Tar was determined using the Rebelein method (Rebelein, 1973). In the late autumn, after the harvest, at each sampling site ten replicate soil cores (diameter = 1.91 cm, depth = 0-30 cm) were randomly collected, and mixed, leading to one homogenised sample per location. The soil sampling period was planned to reduce as much as possible the effects of fertilising supply, the latter being usually supplied in early spring. Soil samples were air dried, grounded to pass a <2mm sieve, and then analysed in duplicate according to the Italian Official Methods for Soil Analysis (MIPAF, 2000) for the determination of soil texture, with reference to the following particle size classes: coarse sand (size limits 2.0-0.2 mm), fine sand (0.02-0.05 mm), coarse silt (0.05-0.02 mm), fine silt (0.02-0.002 mm), clay (< 0.002 mm).

## 3.3 Results and discussion

Table 1 shows the principal statistical parameters of PLS for compositional data. In this case, with two components, the model assures a good value of X and Y variation. Moreover the predictive ability for the first components measured by the Q2 is 0.686 and for the second is 0.116.

| **Num. Com.** | **R2X** | **Eig** | **R2Y** | **Q2** |
|---|---|---|---|---|
| 1 | 0.567 | 1.48 | 0.705 | 0.686 |
| 2 | 0.212 | 1.67 | 0.064 | 0.116 |

Table 1: Statistical parameters: R2X is the fraction of Sum of Squares (SS) of the entire X's explained, Eig is the eigenvalue, R2Y is the fraction of Sum of Squares (SS) of the entire Y's explained and Q2 is the fraction of the total variation of the Y's that can be predicted.

The results of the compositional analysis (Fig. 1), highlight, as expected, significant relationships among the considered grape variables and demonstrate the influence of soil texture within the specific pedo-enviornmental conditions of the Telesina Valley vineyard district.

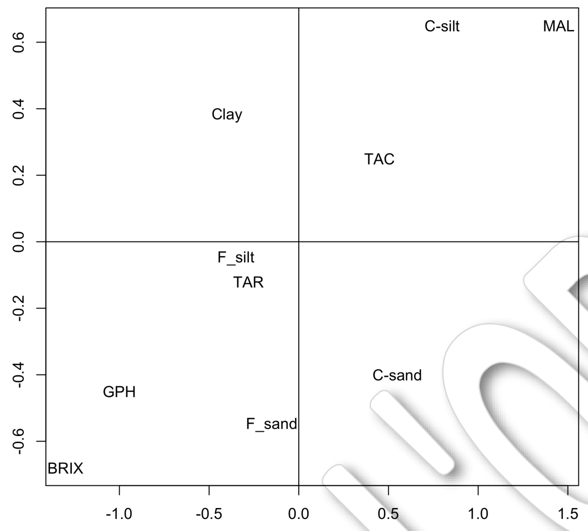


Figure 1: Loading scatter plot.

The projection of Malic acid (Mal) and total soluble solids (Brix) either on the first or second axis are far from the centre of axes and positioned in the opposite directions, meaning that a negative relationship exist between the above variables. Negative relationships between Mal and Brix are known from the literature. Malic acid, along with Tar, slowly increases in concentration until *veraision*. After *veraison* the concentration of Mal rapidly declines, as a consequence of the reduced synthesis, as well as of the metabolic utilisation in replacing glucose as the major respiration substrate, particularly during the later stages of ripening. Conversely, as acidity falls during ripening process, sugar levels (here express as Brix degrees) increase, mainly as a consequence of increasing synthesis of glucose and fructose, which represent the principal grape sugars (Jackson, 2000).

Malic is also positively related to titratable acidity (Tac). As known (Fregoni, 1999, Jackson, 2000), malic and tartatic acids (Tar) are, by far, the most important acids found in grapes. As referred previously, after *veraison* the concentration of Mal rapidly decline, as a consequence of the reduced synthesis. Different from Mal, Tar content usually tends to stabilise after *veraison* (Jackson, 1994). This acid declines slowly and less intensely than Mal during ripening. Thus the decrease of TAc, during ripening, is nearly related to the degradation of malic acid. The positive relationship between Mal and TAc explain also the negative relationships between each of the above variables and grape pH (GpH). According to Figure 1, Tar is close to the centre of axes, thus it means that some information may by carried over on other axes, then any interpretation pertaining this variable might be hazardous. Pertaining the relationships between soil texture and grape composition, it can be observed (Fig. 1) that 1) coarse and fine sand are positively correlated with Brix (particularly) and GpH, and negatively correlated with Mal (particularly) and TAc; b) conversely, coarse silt and clay are negatively correlated with the above grape variables, and positively with Mal (particularly) and TAc. All the above relationships are explained by the influence of texture on soil temperature and water content. In particular, coarser textured soils (i.e. dominated by the sand fraction), who have a better thermal conductivity, lead to a good root activity during the vegetative phase and to the reduction of reduction of vegetative activity during summer time, when soil profile becomes dry, thus allowing a good accumulation of sugars, as well as of aromas and polyphenols in berries (Fregoni, 1999). Finer textured soils (i.e., soil dominated by clay and silt fractions) tend to have higher water content than coarser textured soils. Water content affects the absorption of nitrogen by vine (Fregoni, 1999), besides than dilution phenomenon in the berry juice. It is recognised that as soil water content increases, nitrogen uptake also increases. Increasing nitrogen uptake produces increasing vine vegetative growth and, subsequently, the reduction of both acids degradation, particularly of Mal and sugar accumulation.

## Bibliography


Aitchison, J. (1982). The statistical analysis of compositional data (with discussion). *J. R. Stat. Soc*. Ser. B,  44, 139–177.

Aitchison, J. (1986). *The Statistical Analysis of Compositional Data.* Chapman and Hall: London, UK.

Boehringer, M., 1980: Food analysis. Boehringer Biochemia Robin, Milano

Brady, N.C.; Weil, R.R.; 2002: The nature and properties of soils. Printice Hall, Upper Saddle River, New Jersey, pp. 960

Fregoni, M.; 1999: Viticoltura di qualità. Stampa Grafiche Lama, Piacenza, pp. 705

Gallo, M. (2003). Partial Least Squares for Compositional Data: an approach based on the splines. *Italian Journal of Applied Statistics*, 15, 3, 349-358.

Hinkle, J., Rayens, W. (1995) Partial least squares and compositional data: problems and alternatives. *Chemometrics and Intelligent Laboratory Systems*, **30**, 159–172.

Jackson, R.S.; 2000, Wine science. Principles and applications. Elsevier Science & Technology Books, pp640

Moio, L.; Romano, S.; Cirella, A.; Fuschino, G.; 1999: Influenza di alcuni fattori viticoli ed enologici sulla qualità del vino rosso prodotto dall'uva Aglianico di Taurasi. Vignevini, 1-2: 80.92

Monaco, A.: Storia di un vitigno. In: Falanghina (Monaco, A, Mustilli, A.C. and Pignataro L., eds). Franco Di Mauro Editore, pp 29-49 (2004).

Morlat, R., Bodin, F.: Characterization of viticultural terroirs using a simple field model based on soil depth. II. Validation of the grape yield and berry quality in the Anjou vineyard (France).

Rayens, W. (2000) *The art of maximizing covariance*. University of Kentucky Technical Report 383.

Rebelein, H., 1973: Verbrafen zur genauen serien Massingen Bestimmunq der Wein und Nilchsaure in Wein und ahnichen Getraken. Chem. Mikrobiol. Technol. Lebnsm. 2, 23

Van Leeuwen, C. and Seguin, G.: The concept of terroir in viticulture. J. Wine Res., vol. 17, n. 1, 1-10. (2006)

White, R.E., 2003: Soil for fine wines. Oxford University Press, New York.

Wold, H. (1966). Estimation of principal components and related models by iterative least squares. In P.R. Krishnaiaah (Ed.). Multivariate Analysis. (pp.391-420) New York: Academic Press.

Wold, S, Sjöström, M., Eriksson, L. (2001). PLS-regression: a basic tool of chemometrics. *Chemometrics and Intelligent Laboratory Systems*, 58, 109–130.

# Appendix

$$\underset{\substack{\mathbf{a}\in\Re^J;\ \mathbf{b}\in\Re^K\\ \mathbf{a}^t\Sigma_Z\mathbf{A}=\hat{\mathbf{0}}\\ \mathbf{b}^t\Sigma_V\mathbf{B}=\hat{\mathbf{0}}\\ \mathbf{a}^t\hat{\mathbf{1}}_J=0;\ \mathbf{b}^t\hat{\mathbf{1}}_K=0}}{\arg\max}\left\{\frac{\operatorname{cov}^2(\mathbf{Za},\mathbf{Vb})}{(\mathbf{a}^t\mathbf{a})(\mathbf{b}^t\mathbf{b})}\right\}=\underset{\substack{\mathbf{a}\in\Re^J\\ \mathbf{a}^t\Sigma_Z\mathbf{A}=\hat{\mathbf{0}}\\ \mathbf{a}^t\hat{\mathbf{1}}_J=0}}{\arg\max}\left\{\frac{1}{(\mathbf{a}^t\mathbf{a})}\underset{\substack{\mathbf{b}\in\Re^K\\ \mathbf{b}^t\Sigma_V\mathbf{B}=\hat{\mathbf{0}}\\ \mathbf{b}^t\hat{\mathbf{1}}_K=0}}{\arg\max}\left\{\frac{\operatorname{cov}^2(\mathbf{Za},\mathbf{Vb})}{(\mathbf{b}^t\mathbf{b})}\right\}\right\}$$

$$=\underset{\substack{\mathbf{a}\in\Re^J\\ \mathbf{a}^t\Sigma_Z\mathbf{A}=\hat{\mathbf{0}}\\ \mathbf{a}^t\hat{\mathbf{1}}_J=0}}{\arg\max}\left\{\frac{1}{(\mathbf{a}^t\mathbf{a})}\underset{\substack{\mathbf{b}\in\Re^K\\ \mathbf{b}^t\Sigma_V\mathbf{B}=\hat{\mathbf{0}}\\ \mathbf{b}^t\hat{\mathbf{1}}_K=0}}{\arg\max}\left\{\frac{\mathbf{b}^t\Sigma_{VZ}\mathbf{a}\mathbf{a}^t\Sigma_{ZV}\mathbf{b}}{(\mathbf{b}^t\mathbf{b})}\right\}\right\}$$

$$=\underset{\substack{\mathbf{a}\in\Re^J\\ \mathbf{a}^t\Sigma_Z\mathbf{A}=\hat{\mathbf{0}}\\ \mathbf{a}^t\hat{\mathbf{1}}_J=0}}{\arg\max}\left\{\frac{1}{(\mathbf{a}^t\mathbf{a})}\underset{\substack{\mathbf{b}\in\Re^K\\ \mathbf{b}^t\Sigma_V\mathbf{B}=\hat{\mathbf{0}}\\ \mathbf{b}^t\hat{\mathbf{1}}_K=0}}{\arg\max}\left\{\frac{\mathbf{b}^t\left(\mathbf{I}-\mathbf{P}_h^V\right)\Sigma_{VZ}\mathbf{a}\mathbf{a}^t\Sigma_{ZV}\left(\mathbf{I}-\mathbf{P}_h^V\right)\mathbf{b}}{(\mathbf{b}^t\mathbf{b})}\right\}\right\}$$

where $\mathbf{P}_h^V=\left(\Sigma_V\mathbf{B}\left(\mathbf{B}^t\Sigma_V\Sigma_V\mathbf{B}\right)^{-1}\mathbf{B}^t\Sigma_V\right)$ and $\mathbf{b}^t\left(\mathbf{I}-\mathbf{P}_h^V\right)=\mathbf{b}^t$ the inner supremum is a principal components problem. The maximum is $\mathbf{a}^t\Sigma_{ZV}\left(\mathbf{I}-\mathbf{P}_h^V\right)\Sigma_{VZ}\mathbf{a}$ and $\mathbf{b}=\left(\mathbf{I}-\mathbf{P}_h^V\right)\Sigma_{VZ}\mathbf{a}$.

So

$$\underset{\substack{\mathbf{a}\in\Re^J\\ \mathbf{a}^t\Sigma_Z\mathbf{A}=\hat{\mathbf{0}}\\ \mathbf{a}^t\hat{\mathbf{1}}_J=0}}{\arg\max}\left\{\frac{1}{\left(\mathbf{a}^t\mathbf{a}\right)}\underset{\substack{\mathbf{b}\in\Re^K\\ \mathbf{b}^t\Sigma_V\mathbf{B}=\hat{\mathbf{0}}\\ \mathbf{b}^t\hat{\mathbf{1}}_K=0}}{\arg\max}\left\{\frac{\mathbf{b}^t\left(\mathbf{I}-\mathbf{P}_h^V\right)\Sigma_{VZ}\mathbf{a}\mathbf{a}^t\Sigma_{ZV}\left(\mathbf{I}-\mathbf{P}_h^V\right)\mathbf{b}}{\left(\mathbf{b}^t\mathbf{b}\right)}\right\}\right\}=\underset{\substack{\mathbf{a}\in\Re^J\\ \mathbf{a}^t\Sigma_Z\mathbf{A}=\hat{\mathbf{0}}\\ \mathbf{a}^t\hat{\mathbf{1}}_J=0}}{\arg\max}\left\{\frac{\mathbf{a}^t\Sigma_{ZV}\left(\mathbf{I}-\mathbf{P}_h^V\right)\Sigma_{VZ}\mathbf{a}}{\left(\mathbf{a}^t\mathbf{a}\right)}\right\}=$$

$$\begin{aligned}
&=\underset{\substack{\mathbf{a}\in\Re^J\\ \mathbf{a}^t\Sigma_Z\mathbf{A}=\hat{\mathbf{0}}\\ \mathbf{a}^t\hat{\mathbf{1}}_J=0}}{\arg\max}\left\{\frac{\mathbf{a}^t\left(\mathbf{I}-\mathbf{P}_h^Z\right)\Sigma_{ZV}\left(\mathbf{I}-\mathbf{P}_h^V\right)\Sigma_{VZ}\left(\mathbf{I}-\mathbf{P}_h^Z\right)\mathbf{a}}{\left(\mathbf{a}^t\mathbf{a}\right)}\right\}\\
&\leq\underset{\mathbf{a}\in\Re^J}{\arg\max}\left\{\frac{\mathbf{a}^t\left(\mathbf{I}-\mathbf{P}_h^Z\right)\Sigma_{ZV}\left(\mathbf{I}-\mathbf{P}_h^V\right)\Sigma_{VZ}\left(\mathbf{I}-\mathbf{P}_h^Z\right)\mathbf{a}}{\left(\mathbf{a}^t\mathbf{a}\right)}\right\}\\
&=\lambda_{h+1}
\end{aligned}$$

where $\mathbf{P}_h^Z=\left(\Sigma_Z\mathbf{A}\left(\mathbf{A}^t\Sigma_Z\Sigma_Z\mathbf{A}\right)^{-1}\mathbf{A}^t\Sigma_Z\right)$ and $\Sigma$. Hence $\left(\mathbf{I}-\mathbf{P}_h^Z\right)\Sigma_{ZV}\left(\mathbf{I}-\mathbf{P}_h^V\right)\Sigma_{VZ}\left(\mathbf{I}-\mathbf{P}_h^Z\right)$ is the largest eigenvalue and $\mathbf{a}_{h+1}$ is the corresponding eigenvector. This is the solution sought if there are satisfy the constrain $\mathbf{a}^t\Sigma_Z\mathbf{A}=\hat{\mathbf{0}}$, $\mathbf{a}^t\hat{\mathbf{1}}_J=0$, and $\mathbf{b}^t\Sigma_V\mathbf{B}=\hat{\mathbf{0}}$, $\mathbf{b}^t\hat{\mathbf{1}}_K=0$:

$$\begin{aligned}
\mathbf{a}^t\Sigma_Z\mathbf{A}&=\frac{1}{\lambda_{h+1}}\mathbf{a}_{h+1}^t\left(\mathbf{I}-\mathbf{P}_h^Z\right)\Sigma_{ZV}\left(\mathbf{I}-\mathbf{P}_h^V\right)\Sigma_{VZ}\left(\mathbf{I}-\mathbf{P}_h^Z\right)\Sigma_Z\mathbf{A}\\
&=\frac{1}{\lambda_{h+1}}\mathbf{a}_{h+1}^t\left(\mathbf{I}-\mathbf{P}_h^Z\right)\Sigma_{ZV}\left(\mathbf{I}-\mathbf{P}_h^V\right)\Sigma_{VZ}\hat{\mathbf{0}}\\
&=\hat{\mathbf{0}}
\end{aligned}$$

$$\mathbf{b}^t\Sigma_V\mathbf{B}=\mathbf{a}^t\Sigma_{ZV}\left(\mathbf{I}-\mathbf{P}_h^V\right)\Sigma_V\mathbf{B}=\hat{\mathbf{0}}$$

$$\mathbf{a}_{h+1}^t\hat{\mathbf{1}}_J=\frac{1}{\lambda_{h+1}}\mathbf{a}_{h+1}^t\left(\mathbf{I}-\mathbf{P}_h^Z\right)\Sigma_{ZV}\left(\mathbf{I}-\mathbf{P}_h^V\right)\Sigma_{VZ}\left(\mathbf{I}-\mathbf{P}_h^Z\right)\hat{\mathbf{1}}_J=0$$

$$\mathbf{b}^t\hat{\mathbf{1}}_K=\mathbf{a}^t\Sigma_{ZV}\left(\mathbf{I}-\mathbf{P}_h^V\right)\hat{\mathbf{1}}_K=0$$