The Saga of PLS

Gaston Sanchez

Contents

Preface

A saga is a story of heroic events and achievements of a personage or family, typically written in medieval Icelandic or Old Norse in the form of a prose narrative. According to Wikipedia, the term saga originates from the Norse "saga" and refers to "what is said" or "story, tale, history." The closest term in English to the word saga is "saw" (as in *old saying*). In this regard, I'm using the word Saga to play with its meaning, and convey the ideas of history, story, and narrative of events associated to Partial Least Squares methods.

Acknowledgements

I've been able to gather invaluable unpublished information and details from my own involvement around the PLS community, as well as from personal communications with leading *PLSers* (via email, skype, and personal meetings).

My heartfelt thanks to Michel Tenenhaus, Wynne Chin, and Vincenzo Esposito-Vinzi. They all have being very gentle to answer to my emails full of questions and inquiries, providing responses full of details and milimetric precision.

Also many thanks to Christian Ringle, José Luis Roldán, and Aida Eslami for their emails with useful information and enriching comments. Likewise, I would like to thank the various friends, colleagues, and, for the lack of a better term, *followers* around the world that have been helping me proofread the content, detecting bugs, and fixing typos. I take full responsibility for any prevailing errata in the text.

Last but not least, you wouldn't be reading this book if it wasn't for the

patience and support of my loving wife Jessica. Not only she was willing once again to be my household editor, but she never complained my occupying of the dining table, and taking over the living room as my personal workspace. If you find any value from the content in this book, which I am sure you will, you owe her almost as much as you owe me.

Berkeley, California.

April 2015.

Kind Words

> "Wonderful, beautiful text on PLS. I was, of course, pleased and flattered over the kind treatment of my father Herman and myself, and most impressed of the great coverage of all essential aspects of PLS, both path-PLS and PLS-regression."
>
> *Svante Wold*

> "Greatly enjoyed The Saga of PLS."
>
> *Nouna Kettaneh*

> "It's a wonderful work."
>
> *Michel Tenenhaus*

> "A fantastic reading... And Gaston is right, I don't remember neither the question he asked me, nor the answer I gave him when he asked me about the PLS regressions within the PLS-PM algorithm."
>
> *Tomàs Aluja*

> "Gaston has done a great job by writing this extremly interesting and useful book."
>
> *Aida Eslami*

"This is a magnificent work, clearly protraying the history and evolution of PLS."

José Luis Roldán

"I just finished my first reading of The Saga of PLS, and I love it."

Diogenes de Souza Bido

"This book brings fond memories of Herman Wold and J-B. Lohmoller who often visited the campus of Case Western Reserve University, Cleveland, while giving talks/seminars and enjoying musical performances by the Cleveland Orchestra."

Nguyen Quan

"I read and reread The Saga of PLS. I've shared it with many members of the PLS community and it was incredibly well received by faculty around the world."

Soumya Ray

"I thoroughly enjoyed reading The Saga of PLS… a very nice introduction to the rich history of PLS."

Derek Beaton

"An admirable work, providing us with outstanding material about PLS, so hard to find elsewhere."

Pablo Cáceres Serrano

"The Saga of PLS was a very good read."

Age Johnsen

Chapter 1

Introduction

The main motivating trigger behind this book has been my long standing obsession to understand the historical development of Partial Least Squares methods in order to find the who's, why's, what's, when's, and how's. It is the result of an intermittent 10 year quest, tracking bits and pieces of information in order to assemble the story of such methods. Moreover, this text is my third iteration on the subject, following two of my previous works: chapter 2 *"Historical Review"* of my PhD thesis (Sanchez, 2009), and the appendix *"A Historical Overview of PLS-PM"* from my book *PLS Path Modeling with R* (Sanchez, 2013).

This is NOT a technical book. It doesn't cover theory, methodological aspects, nor technical details of how the various PLS methods work (no discussions about algebra, computational steps, interpretation issues, etc.). This is also not a book written with a particular reader in mind. Instead, I've written its content to organize the vast material I've collected so far, which includes not only information from papers, chapters, proceedings, and books, but also thoughts, memories, analysis, interpretations, and personal opinions. Having said that, I imagine that this book can be used as a **companion reading** for any course, workshop, or seminar about PLS methods, expecting to be enjoyed by anyone interested on this topic. Regardless of whether you're just taking your first steps on the PLS arena, or you've already traveled a long way around PLS territory, I'm sure you'll find some value in the content of this work.

By writing this book, my aim is to shed some light about the teaching and understanding of the historical background surrounding Partial Least Squares methods. I'm convinced that this material will provide a fresh perspective and create awareness, debunk myths, and clarify misunderstandings. I'm not the one to judge whether the story of PLS has elements of an heroic tale. But I can tell you for sure that this story is formed with a long and complicated series of events, having scandinavian authors as protagonists. Hence the meaning of the title "The Saga of PLS."

1.1 The proof of the pudding

In his classic *"Soft Modeling: The Basic Design and Some Extensions,"* Herman Wold includes the following footnote on page 24 (Wold, 1982b):

> "Most nonlinear iterative techniques of estimation are lacking an analytic proof of convergence. The proof of the pudding is in the eating."

Referring to the lack of a proof of convergence for his algorithm—when applied to the general case of multiple relations among multiple sets of variables—, Wold turned to the old saying of the **proof of the pudding** to invite readers to try for themselves his methodology. Michel Tenenhaus (pronounced *Tenen-os*), a French statistician and world renowned PLS expert, has also used the same famous saying when presenting Partial Least Squares methods to new audiences. For many years, Michel has included William Camden's (1623) version "All the proofs of a pudding are in the eating, not in the cooking," in the last slide of most of his presentations about PLS, inviting users to *taste* the method and experience for themselves whether they like it or not.

For almost three decades, the convergence of Herman's main algorithm was an unresolved issue. This was due in part to the various options in which the algorithm can be specified. Depending on the input settings, computations converge most of the times. But there are a few exceptions. Today, we have a much better understanding of the mechanics of the algorithm, and it is possible to prove the convergence under certain circumstances. Unlike

Wold, we also have a better comprehension of how the steps operate, and we can even tweak the algorithm in a such a way that we are certain of what's going on inside of it. Thanks to these advancements, a couple of years ago, Tenenhaus slightly modified the pudding phrase to:

> All the proofs of a pudding are in the eating… but it will taste even better if you know the cooking.

He's right. Let me add to it that it doesn't hurt to know about the story and history behind the recipes, their chefs, their cooking styles, and their tasting preferences. If anything, it will make your tasting of the pudding irresistible.

1.2 Outline

This book is organized into three major sections:

1. Introductory Review (chapters 2 - 4)
2. Historical Narrative (chapters 5 - 9)
3. Assessment and Conclusions (chapters 10 - 11)

Part 1 covers a general introduction describing general aspects about Partial Least Squares methods. Part 2 presents a historical narrative focused on Herman Wold and the series of events that led him to develop the so-called *PLS Soft Modeling* framework. It also contains a narrative of the role played by Svante Wold and the development of PLS Regression. Instead of merely listing various events and polluting the text with dozens of citations, I've preferred to use a more fluid narrative style. Finally, the third part provides remarks and critical assessments of several points often ignored in the PLS related literature. My purpose is to give you a different perspective of Partial Least Squares that, hopefully, will help you gain insight into a better understanding and comprehension on the subject.

Chapter 2

You've got to be kidding me

The first time I heard about PLS was in October 2004 when I first met professor Tomàs Aluja—my former PhD adviser. I had just recently moved to Barcelona to start my graduate studies in statistics, and I had been designated Tomàs's new mentee. At the end of our first meeting when I was about to leave his office, Tomàs asked me a seemingly intrascendental question: "Have you ever heard about PLS?"

"P. L. S.?"

I repeated those letters to myself without pronouncing any words. Fearing to give a bad impression to my brand new mentor, I quickly scanned the hard-disk of my brain in search for any hint. Unsuccessfully struggling to find any answer, I ended up with a perfect clueless expression in my face.

I had no idea what PLS was or what it meant.

The reason why Tomàs asked me about PLS was because he was in charge of organizing *PLS'05 - The 4th International Symposium on Partial Least Squares and Related Methods* for the following year (September, 2005). He wanted to know if I would be willing to help him with some of the organizational details for the symposium. Little did I know the tremendous impact those three mysterious letters would have in the coming years of my analytics career.

I became involved with PLS in a very unintended way. One of my first duties as part of the *PLS'05* staff, was being in charge of receiving the abstracts and paper submissions for the symposium, making sure they complied with the strict guidelines-for-authors. I spent my time sending emails and pestering authors kindly requesting them to fix the page layout, change the font size, reduce the number of pages, and things like that. One of the orders Tomàs gave me was to read all the papers. Grudgingly, I followed his command and sucked it all up. However, *reading* doesn't necessarily mean *understanding*. From all the submitted papers, I was able to understand three of them… at the most. And even then my understanding was very superficial. The common thing about those papers was their very applied and digestable examples that prevented me from getting lost with all the unintelligible details. By the way, this is NOT how anyone should ever try to learn about PLS methods.

I had a more formal and intensive introduction to the PLS framework with Tomàs's course "Advanced Methods for Multivariate Data Analysis." The outline of the course was based on the great book *La Régression PLS: Théorie et Pratique* by Michel (Tenenhaus, 1998). The book is a marvelous piece, with Michel's unique eloquent writing style, explaining things in a very detailed way, using a clear notation, and full of examples and graphics. I guess everything would have been fine if it wasn't for the fact that I had a very limited knowledge of French. Which it only makes it harder when you're not only a PLS novice but also a French-language beginner. Luckily, *google translate* was already available at that time, providing a considerable help in deciphering some of the book's content.

Tomàs's class was a very hands-on course. Each student was responsible of studying a chapter from the book, programming the corresponding method in R, and giving a lecture to the rest of the class. I got to present chapter 8—PLS Regression for one response variable. Tomàs kept for himself the presentation of the PLS Path Modeling (PLS-PM) approach for the last lecture. That day he introduced the technique with an example using the famous Russett dataset, and then he proceeded to talk about how the method worked. When he got to the description of the PLS-PM algorithm, however, I was having a hard time understanding all the steps. The worst part was my fruitless attempt trying to connect the PLS regression algorithms with the PLS-PM algorithm. The question that was driving me crazy was: Where on earth were the PLS regressions in the PLS-PM algorithm? I raised my hand and asked Tomàs if he could please show the rest of the class where the PLS regressions were. To my surprise, his answer was even more disconcerting: "Regressions? What regressions? This has nothing to do with PLS Regression. This is PLS Path Modeling." I'm pretty sure Tomàs does not even remember these details. But they definitely had a profound impact on me.

All I remember thinking was something like: "You've got to be kidding me, this PLS stuff makes no sense whatsoever!" I was definitely biting off more than I was able to chew.

I won't lie to you. My introduction to the PLS world was not a smooth one. Besides my very limited knowledge of French, another difficulty was my equally limited knowledge of the techniques in Michel's book like Canonical Correlation Analysis, Redundancy Analysis, and Inter-Battery Factor Analysis. One of the morals from this experience was that I needed to learn French; the other moral was that this was also not the best way to start learning about PLS from scratch.

I felt so confused and disoriented, but eventually I managed to my find way out through the PLS labyrinth over the following years. My biggest consolation, to my suprise, was when I heard Michel Tenenhaus story in front of the audience at one of the PLS Research Workshops in Paris, 2009. He told us about his PLS tale: having started studying PLS regression, he was later referred to PLS Path Modeling, only to find out that, initially, nothing made sense to him. I was so relieved in knowing that I was not the only one with the same big confussion. Even a PLS *jedi master* like Michel had faced the same puzzle I had struggled with.

I became obsessed with the story behind PLS since the very first moment I started to work on my PhD dissertation about PLS Path Modeling. As a grad student in the fall of 2005, I remember googling about "Partial Least Squares" and getting links back that talked about PLS Regression. When I tried to be more specific and searched for "Partial Least Squares Path Modeling," there was nothing but again PLS Regression retrieved results only. Why was there almost nothing about PLS Path Modeling? There were virtually no resources about the original works by Herman Wold, or the so-called Structural Equation Models via PLS approach. Moreover, most of the available resources on the Web were basically chemometrics-related material. It is amazing to see how much things have changed in just a decade. But it hasn't been easy. It has required a titanic effort coordinating the work of a large number of scholars. Gradually, the barriers have been pushed in order to revive a methodological framework that was heading toward the annals of statistical methods that sadly end up in the dead archives of university libraries.

One of the very first things I knew I wanted to do was to find out more about the history of PLS methods. The main resource I had in my hands was a tutorial on PLS Path Modeling, which was basically a draft for the famous article *"PLS Path Modeling"* (Tenenhaus et al., 2005). I started tracking the references listed in the tutorial, and I was happy to find that the Economics and Math libraries of the University of Barcelona had copies of the two volumes *Systems under indirect observation: Causality, structure, prediction* (Wold, 1982b), the intimidating treatise *Latent Variable Path Modeling with Partial Least Squares* (Lohmöller, 1989), the *Encyclopedia of Statistical Sciences* with the entry of PLS (Wold, 1985), and *The Making of Statisticians* containing the autobiographical essay of Herman (Wold, 1982a). Little by little I started to dig out and uncover fragments of a fascinating story.

All I wanted was to be able to answer the *who*, *when*, *how*, and *why*. Who created PLS? When was it developed? How was the development process? And why? How does something such as PLS is developed? What were the circumstances that gave birth to that framework? In this sense, the quest became a quest not only about PLS itself but also about its creator, about his mind, his ideas, and surrounding context. Among all my inquires, one question has remained for years in the back burner of my head: Why was PLS Path Modeling not that well known? Think about it. By 2005, the first works about PLS had already been present for almost 30 years! Yet, having access to those references was—and still is—ridiculously difficult. I think I have now most of the elements to answer the questions that have been lingering since that fall of 2005.

Chapter 3

A little bit about Partial Least Squares

One of the ever present tasks of almost every data analysis endeavor is the study of relationships between variables. In its simplest form, the analyst wants to know whether two values are related with one another. For instance, I might be interested in studying if there is a relation between the number of calories in the coffees my friends and I drink, and the price we pay for them. The question about the relation between calories and price can be further broken down in other inquiries: What is the nature of the potential relationship? What is its magnitude? Or how important is it? I could also expand the analysis and ask about a possible dependency: Can I predict the prices of coffees based on the number of calories they contain? Or viceversa: Knowing the price of a coffee, what can I say about its number of calories? To answer these questions we can turn to the rich set of analytical tools that statistical methods give us by using a number of graphical displays and charts, measures of simultaneous variation such as covariance, coefficients of association such as correlation, and modeling options such as regression analysis.

Often, we don't have just two variables but many more. If we were hired to do a deeper and broader analysis on data from coffee shops, we surely would have more available information like the area size of the establishments, or records about frequent customers. Hence we could ask about the relationship between the number of square meters and the cost of the coffees. Furthermore, we can think about the differet aspects that go into the preparation of a cup of coffee—the beans, water, milk, sugar, etc—and how they may be associated with the financial performance of the businesses (earnings, losses, costs, etc). These questions and many others arise naturally when we do data analysis, and we aim to tackle them at various levels of statistical complexity.

Depending on the circumstances, we may face situations in which the analyzed data are divided—or can be divided—into sets of variables. With the hypothetical coffee shops data, one set of variables may be comprised by quantity of ingredients (beans, water, milk, sugar), a second set may be formed by utilities consumption (gas, electricity, water, internet, etc), and a third set may be composed of sales by size of container (sold units by size: small, medium, large). Sometimes we just care about describing and summarizing the relationships. This is typically done in the initial phase of the analysis when we are in exploratory mode. Sometimes we may be interested in relationships of prediction and explanatory nature. Based on the analysts's preconceived ideas or hypotheses about the data, more advanced and sophisticated tools can de devised. In all these cases we can rely on a wide range of multivariate statistical methods that allow us to generalize the analysis of two variables into two or more sets of variables. Among such methods, Partial Least Squares techniques stand out as one remarkable toolbox that, simply put, provides a versatile data modeling platform for analyzing relations among one or more sets of variables.

3.1 PLS

Rooted in a couple of data model-building and computational ideas developed by Herman Wold in the 1960s, **Partial Least Squares** methods have traveled a long journey until reaching the high praise and recognition they receive today. Experiencing gradual changes, and bifurcating into several branches, PLS methods have gone through a fascinating—although not always smooth—evolutionary process. Nevertheless, they all have proven to be tremendously valuable on both theoretical and practical aspects.

At its heart, Partial Least Squares provide a versatile platform for analyzing multiple relationships among one or more sets of variables—measured on some objects. Among their attractiveness we can highlight several features:

- their deceptively simple iterative mechanisms
- their ease of programming implementation
- their estimation ideas anchored on the principle of least squares
- their marked geometrical flavor of projection-based methods
- their inherent dimension reduction nature
- their rich possibilities for data visualization
- their strong inductive spirit
- their fit within statistical learning approaches

Above everything else, the most captivating trait—in my opinion—is that Partial Least Squares offers us a wide analytical platform that covers a large number of multivariate data analysis techniques. Plus, it gives us the ability to connect a number of seemingly unrelated methods. Such a platform, encompassing both its computational side, and its flexible analytical essence, is of a great richness and applicability very few times encountered in other data analysis approaches.

3.1.1 A broad definition of PLS

If what you are looking for is an official definition of PLS I'm afraid you won't find such thing. Instead, we could give different meanings to the PLS term. From a narrow point of view, we could talk about PLS as an algorithmic template for computing several multivariate models. From a more applied angle we could also talk about PLS as a family of regression-type data analysis methods. From a richer and more comprehensive standpoint, we can regard PLS as an umbrella term for dealing with multirrelational systems of one or more sets of variables following the principle of partial least squares. It is the latter description that I like to consider for defining PLS as a general framework. Here's my definition. Broadly speaking:

> Partial Least Squares is a versatile multivariate data modeling framework for analyzing multiple relationships among one or more sets of variables measured on some objects.

One thing for sure is that there's a large number of methods labeled with the PLS acronym. Such an amazing variety of methods and approaches, so diverse and plural, makes them seem practically impossible to view under a united scope. In my opinion, many headaches would be avoided if there was a unique reason for a method to be labeled with the PLS acronym. For better or worse, there is not just one but several sources that make a method be sheltered under the PLS brand. Methods carrying PLS as prefix or suffix are so varied and so imbued in different fields and areas of application that it seems hard to picture them all under the same lens.

The literature about PLS methods is incredibly rich. If you take a look at the dozens of references about Partial Least Squares, you will be shocked with a mind blowing puzzle. The major benefit of such abundance of material is the r ichness of concepts, principles, applications, extensions, and new proposals. The ugly side is that there's no universal view, no unified notation, no unique terminology, and even worse, slightly different connotations of the term PLS. Which only makes it easy to get lost along the way. After all, this reflects the evolution of PLS methods, how they have spread out, how they have permeated into different subfields, and how they have come to occupy multiple places within the collective consciousness of Multivariate Data Analysis. Everywhere a PLS method has taken up residence, it dresses in different clothes, does its hair differently, and speaks a slighty different dialect.

In summary, as soon as you take a first step into PLS territory, you can immediately get disoriented in the middle of the trees without being able to appreciate the forest in all of its magnitude. No wonder why many practitioners are confused about exactly what PLS is, what it does, how it works, and how they can benefit from it.

3.2 Some words of caution

I would be a fool to tell you that there is one single universal perception of Partial Least Squares. Depending on who you talk to, you can get different meanings for the term PLS. If you ask your favorite search engine about "PLS," you might get answers like "please," or "political science," which is obviously not what we are talking about. If you search for the term "partial least squares," I bet you will find most results related to what is known as *PLS Regression*—the most popular version of PLS methods. Not surprisingly, most people think of PLS as just another technique for solving regression-like problems. This is not a mere coincidence but the result of the intricated history behind PLS, splattered with random twists of fate. Now, while it is true that the regression version may be the most common presentation people know, it is by no means the only one. For other people, PLS is an approach for estimating what are known as Structural Equation Models with latent variables.

3.2.1 Path Modelism and Regressionism

Today we can distinguish two main branches of Partial Least Squares approaches: 1) the **Path Modeling** branch, and 2) the **Regression** branch. The former introduced by Herman Wold, and the latter headed by Svante Wold. The recognition of these two large categories has to do with the way in which they have subsequently unfolded. By taking different directions, the branches have produced two major *movements* that for the most part, have grown apart, remained disconnected, and even unaware of one another in some areas of application.

3.2.2 PLS: Noun or Adjective?

The study of Partial Least Squares (PLS) methods begins with potential confusion—the term "PLS" has become broader and looser over the years. One problematic area has to do with the way authors use the term PLS. Sometimes PLS is used as a noun, while others is used as an adjective. As a noun, "PLS" is not one method but a set of methods. Talking about Partial Least Squares as a single methodology is like talking about Athletism as a single discipline, or like talking about the Himalayas as a single mountain. Hence, we should always let users/readers know that PLS implies a big family of methods.

As an adjective, "PLS" describes the *flavor* of a given method. Generally, authors have baptized PLS algorithms after the methodology they sprout from. In this sense, the acronym PLS is used as a label to indicate the estimation approach of a given technique. For instance, PLS as used in "PLS Regression" works as an adjective, describing the estimation procedure for a regression model.

3.2.3 Algorithmic approaches

Another problem is when we hear or read about "the PLS algorithm." The truth is that PLS methods are a bit of an anomaly in that they are both **methodologies** (with specifications, assumptions, and motivational reasons) and **algorithms** (with computational and operational aspects). This means that all PLS methods have their associated algorithm. Often, many different PLS methods are presented by their authors in such a way that the unaware reader is mislead to believe that a particular method is *the* PLS method. This is probably one of the main sources of confusion around the PLS literature. There is not such thing as *the PLS algorithm*, instead there are the PLS algorithms (in plural). For example, there is the PLS regression algorithm, the PLS path modeling algorithm, the PLS algorithm for Principal Components Analysis (a.k.a. NIPALS-PCA), or the PLS algorithm for Canonical Correlation Analysis (a.k.a. NIPALS-CCA), to mention but a few.

If there is anything in common among Partial Least Squares methods is that they all have an associated algorithm with a fairly uniform format. PLS methods proceed in a less intuitive way compared to classical statistical procedures where an optimization criteria is solved algebraically. PLS approaches are not formulated in terms of a global criterion to be optmized. That is, PLS procedures are stated without any maximization or minimization criterion. Just the data decompositions and the system of regression equations are declared. Usually, we express a model (in PLS mindset) in such a way that we identify the components, and the equations between components. Instead of looking to derive an analytical solution, we walk through the solution via a series of repetitive sequential steps until reaching a good stable approximation. In other words, Partial Least Squares are pure algorithmic approaches that consists of a series of steps approximating a stable solution. In many cases, the PLS algorithms coincide with standard algebraic solutions—usually involving an eigendecomposition.

3.2.4 Fragmentation

Another problematic issue is that PLS methods are usually presented in a fragmented way. Most book chapters, papers, courses, workshops, and even conferences deal with only a subset of the dozens of methods, rarely covering a panoramic view. The main reasons for this compartmentalization has to do with historical issues, and the way PLS methods have evolved. Furthermore, accessability to the original references is largely difficult to most practitioners and users.

The initial versions of PLS methods are documented in the 1960s works of Herman Wold, and there is no controversy about the PLS origins. Throughout the entire body of references and publications over several decades about PLS methods in general, all authors acknowledge the roots of the so-called NIPALS procedures in Wold (1966b), Wold (1966a), and Lyttkens and Wold (1969). However, the place PLS methods occupy in contemporary multivariate statitics is a bit more complicated. PLS's rising global popularity has led to disagreements about what constitutes an authentic PLS algorithm. The same name is used to refer to different things, while at the same time different terms are used for naming the same concept. Even within PLS community, ideas about what PLS is varies and has caused debate and tension every once in a while.

3.2.5 The two Wolds

Another source of confusion is the fact that two of the main leading authors behind PLS methods share the same last name "Wold". The unaware reader may think that Wold is one single author without realizing that it can be either Herman (the father) or Svante (the son). In fact, the way and style in which the two Wolds presented their works have left a profound footprint in their ulterior developments. The framework of Herman, emerged from econometric's systems of equations with latent variables, seems to have insurmountable disparities with Svante's Regression Models stemmed from chemometrics. Both branches, with its various subdivisions, show contrasting differences at various levels, notably at the area of application, but also at a philosophical, ideological, conceptual, language, technical, spread and diffusion levels. Albeit their common mathematical and operational elements, and even their shared genes, the contrasting physical appearance between both frameworks puts an illusory divide between them that can easily mislead all inexperienced users, and even some well versed PLS *connoiseurs*.

3.2.6 Pedagogy of PLS

Another hurdle for studying PLS methods has to do with the fact that they have evolved on top of existing techniques, which produces some pedagogical side-effects. Why? Because if you want to learn about a specific PLS technique, usually you must go through a double learning process. Consider for instance Principal Components Analysis (PCA), and its corresponding PLS version NIPALS-PCA. If you want to study NIPALS-PCA, you would begin studying PCA, and then switch to the NIPALS-PCA algorithm. In general, you must first learn about the general problem—and its standard solution—, be it regression, discrimination, principal components, or canonical correlation, just to mention some of them. And then you have to learn how to look at a given technique from the PLS angle. Consequently, this adds an extra layer of concepts, terms, and jargon that you have to deal with: one from the standard approach, and the other one from the PLS approach. No wonder why PLS methods, despite their practical usefulness and methodological attractiveness, have no reserved seats in undergraduate syllabus, and have limited attention in postgraduate programs.

3.2.7 Current Status

Today Partial Least Squares methods are well recognized in pretty much all fields of knowledge where they have been applied. There's a tradition of biannual symposiums exclusively dedicated to discuss advancements and state-of-the-art around PLS and related methods. We also have special publications such as the *Handbook of Partial Least Squares* (Esposito-Vinzi et al., 2010), *New Perspectives in Partial Least Squares and Related Methods* (Abdi et al., 2013). On the computational side a wide range of software programs are available both commercially and free. The largest group of tools is the dozen of R packages freely available in CRAN; there are also XLSTAT plugins (by Addinsoft) for MS Excel; SmartPLS (Ringle et al., 2005), ADANCO, SIMCA (by Umetrics), procedures for SAS, and libraries for Matlab, and Python. Also important is the amazing amount of online learning resources with tutorials, slides, videos, seminars, webinars, workshops, and specialized courses. All this ecosystem reflects the success of PLS methods in industry, academia, research, and governmental spheres.

But it hasn't always been like this.

Bibliography

Abdi, H., Chin, W., Vinzi, V. E., Russolillo, G., and Trinchera, L., editors (2013). *New Perspectives in Partial Least Squares and Related Methods*. Springer.

Esposito-Vinzi, V., Chin, W., Henseler, J., and Wang, H., editors (2010). *Handbook of Partial Least Squares*. Springer.

Lohmöller, J.-B. (1989). *Latent Variable Path Modeling with Partial Least Squares*. Physica-Verlag.

Lyttkens, E. and Wold, H. (1969). Nonlinear iterative partial least squares (nipals) estimation procedures. *Bull. Inst. Internat. Statist.*, 53:29–51.

Ringle, C., Wende, S., and Will, A. (2005). *SmartPLS*. SmartPLS.

Sanchez, G. (2009). *Pathmox Approach: Segmentation Trees in Partial Least Squares Path Modeling*. PhD thesis, Universitat Politecnica de Catalunya.

Sanchez, G. (2013). *PLS Path Modeling with R*. Trowchez Editions, 1st edition.

Tenenhaus, M. (1998). *La Régression PLS. Théorie et Pratique*. Editions Technip, Paris, 1st edition.

Tenenhaus, M., Esposito-Vinzi, V., Chatelin, Y.-M., and Lauro, C. (2005). Pls path modeling. *Computational Statistics & Data Analysis*, 48:159–205.

Wold, H. (1966a). Estimation of principal components and related models by iterative least squares. In Krishnaiah, P. R., editor, *Multivariate Analysis. Proceedings of an International Symposium held in Dayton, Ohio, June 14-19, 1965*, pages 391–420. Academic Press.

Wold, H. (1966b). Nonlinear estimation by iterative least square procedures. In David, F. N., editor, *Research Papers in Statistics. Festschrift for J. Neyman*, pages 411–444. Wiley.

Wold, H. (1982a). Models for knowledge. In Gani, J., editor, *The Making of Statisticians*, pages 190–212. Applied Probability Trust, Academic.

Wold, H. (1982b). Soft modeling: The basic design and some extensions. In Joreskog, K. G. and Wold, H., editors, *Systems under indirect observation: Causality, structure, prediction*, volume 2, pages 1–54. North-Holland.

Wold, H. (1985). Partial least squares. In Kotz, S. and Johnson, N. L., editors, *Encyclopedia of Statistical Sciences*, pages 581–591. Wiley.