

La régression PLS 1, cas particulier de la régression linéaire séquentielle orthogonale (RLSO)

Jacques Goupy

ReConFor, 24 avenue Perrichont. 75016 Paris. France.

jacques@goupy.com

RÉSUMÉ

La régression linéaire classique fournit une seule solution souvent basée sur le critère des moindres carrés. Lorsqu'il y a beaucoup de variables, on risque d'obtenir un modèle surparamétré, c'est-à-dire modélisant les erreurs. Pour éviter cette surparamétrisation, la régression PLS a été introduite car étant une régression séquentielle, elle permet d'arrêter le processus de régression avant de modéliser l'erreur. Mais la PLS ne propose que quelques solutions et ne permet pas vraiment d'optimisation. Nous proposons dans cet article une régression séquentielle orthogonale beaucoup plus souple que la PLS, la régression linéaire séquentielle orthogonale ou RLSO. Cette régression permet d'éviter la surparamétrisation et elle facilite l'optimisation de la modélisation car les directions de projection peuvent être choisies au mieux des intérêts de l'étude. Les solutions possibles proposées par la RLSO sont en nombre infini. Ce qui laisse la possibilité d'utiliser toutes les techniques d'optimisation pour trouver la ou les meilleures solutions.

Comme la régression linéaire séquentielle orthogonale (RLSO) est nouvelle, nous avons développé toutes les étapes de calcul.

Mots clés: Régression séquentielle, Régression PLS1, Régression orthogonale.

ABSTRACT

The classical linear regression provides just one solution often based on the least squares criterion. When there are many variables, one can obtain a model with too many surparameters and consequently a model which modelizes the error. To avoid such an over parameterization, the PLS regression has been introduced. PLS is a sequential regression and it enables to stop the regression process before the error modelling. But the PLS regression has only a few numbers of solutions and the optimization is not really possible. It is why a new type of sequential regression is proposed in this article: the sequential orthogonal linear regression or SOL regression. It is much more versatile than the PLS ; it enables to avoid surparametrization and facilitate the modelling optimization because the directions of projection can be chosen at the best for the study. The number of the solutions proposed by the SOL regression is infinite and all the techniques of optimization can be used.

As the SOL regression is new, all the calculations have been largely explained.

Key words: sequential regression, PLS1 regression, orthogonal regression.