

中国科学技术大学

占位符

Covfefe

Supervised by: USTC

2020 年 10 月 24 日

目录

1	占位符	2
1.1	偏最小二乘回归分析	2

1 占位符

1.1 偏最小二乘回归分析

概述

研究两组或多重相关变量间的相互依赖关系，用自变量 (预测变量) 去预测因变量 (响应变量)。

常用方法：主成分回归分析 (PCR)、最小二乘准则下的经典多元线性回归分析 (MLR)、偏最小二乘 (PLS)。

应用场景

偏最小二乘回归提供一种多对多线性回归建模方法，两组变量个数很多，且都存在多重相关性，样本量少。

例如：身体素质 (身高、体重等) \rightarrow 体能训练结果

算法

经典算法如算法1所示。

Algorithm 1: 偏最小二乘回归分析算法

Data: 自变量集 X ; 因变量集 Y

Result: 达到满意精度的回归方程

- 1 提取 X, Y 的第一对成分 t_1, u_1 ，并使 t_1 与 u_1 变异信息和关联程度最大;
 - 2 **repeat**
 - 3 建立 X, Y 对 t_1 的回归模型;
 - 4 优化残差阵;
 - 5 得到 r 个成分，取前 1 个成分，解得 Y 的偏最小二乘回归方程;
 - 6 交叉有效性检验，要求抽取 h 个成分后， $PRESS(h)/SS(h-1)$ 越小越好;
 - 7 **until** 回归方程精度没有达标;
-

有一种更简单的算法，直接在残差阵提取成分 $t_1 \sim t_r$ ，无需提取 u_h 。