【建模应用】PLS偏最小二乘回归原理与应用

出处：http://www.cnblogs.com/duye/p/9031511.html

1.回归

"回归"一词来源于对父母身高对于子女身高影响的研究。有人对父母的身高与子女身高做统计，发现除了父母高则子女普遍高的常识性结论外，子女的身高总是"趋向"于人类平均身高，最早"回归"一词就来源于此，即子女的身高总是回归于人类平均身高。

现代意义上的回归，是研究因变量对自变量的依赖关系的一种统计分析方法，目的是通过自变量的给定值来估计或预测因变量的均值。它可用于预测、时间序列建模以及发现各种变量之间的因果关系。简单地说，回归就是去分析因变量与自变量之间的关系，从而为分析数据、预测数据提供科学的、合理的方法。

2.回归的方法

目前常用的回归方法有以下几种：

- **多元线性回归**：入门练习常见，但用在模型上基本无用，因为多个变量之间难免有复杂的相关性，多元线性回归不能处理多个自变量之间的"糅合"性。
- **逻辑回归**：当预测的是1/0时使用。 这也是机器学习中的一种常用二分类方法。
- **主成分回归**：类似于主成分分析，将具有相关性的多维变量降维到互不相关的少数几维变量上，可以处理变量之间具有相关性的情况。
- **岭回归**：同上，但方法不同。

- **偏最小二乘回归**：当数据量小，甚至比变量维数还小，而相关性又比较大时使用，这个方法甚至优于主成分回归。

3.回归的检验

建模是最重要的，但好的回归模型是需要检验的，否则你的模型就会显得"苍白无力"。回归结果好与坏，应该怎么检验呢？从以下几方入手：

- **自变量与因变量是否具有预期的关系**。如果有非常不符合逻辑的系数，我们就应该考虑剔除它了。
- **自变量对模型是否有帮助**。如果自变量的系数为零（或非常接近零），我们认为这个自变量对模型没有帮助，统计检验就用来计算系数为零的概率。如果统计检验返回一个小概率值（p值），则表示系数为零的概率很小。如果概率小于0.05，汇总报告上概率（Probability）旁边的一个星号（*）表示相关自变量对

模型非常重要。换句话说，其系数在95%置信度上具有统计显著性。

- **残差是否有空间聚类**。残差在空间上应该是随机分布的，而不应该出现聚类。
- **模型是否出现了倾向**性。如果我们正确的构建了回归分析模型，那么模型的残差会符合完美的正态分布，其图形为钟形曲线。
- **自变量中是否存在冗余**。建模的过程中，应尽量去选择表示各个不同方面的自变量，也就是尽量避免传达相同或相似信息的自变量。 **评估模型的性能**。评估**矫R2值**，有时还要加上修正的Akaike信息准则/Akaike's information criterion (**AIC**)，效果是否好。

偏最小二乘回归

1.偏最小二乘回归的思想：

一般来说，能用主成分分析就能用偏最小二乘。偏最小二乘集成了主成分分析、典型相关分析、线性回归分析的优点。在普通多元线形回归的应用中，我们常受到许多限制。 最典型的问题就是：自变量之间的多重相关性。并且**有的时候样例很少，甚至比变量的维度还少，变量之间又存在多重相关性**。偏最小二乘回归就是为解决这些棘手的问题而生的。

举个例子，比如现在，有一堆因素（X1,X2,...Xn）(这是自变量)，这些因素可以导致（Y1,Y2,...Yn）（这是因变量），给的样例很少，而我们又完全不清楚自变量之间、因变量之间存在的关系，这时问自变量与因变量之间到底是一个什么关系？这就是偏最小二乘要解决的问题。

2.偏最小二乘回归建模原理：

设有 q个因变量{y1,...,yq}和p自变量{x1,...,xp}。为了研究因变量和自变量的统计关系,我们观测了n个样本点,由此构成了自变量与因变量的数据表X={x1,...,xp}和.Y={y1,...,yp}。偏最小二乘回归分别在X与Y中提取出成分t1和u1(也就是说, t1是x1,x2,...,xq的线形组合, u1是y1,y2,...,yp的线形组合).在提取这两个成分时,为了回归分析的需要,有下列两个要求:

(1) t1和u1应尽可能大地携带他们各自数据表中的变异信息;

(2) t1与u1的相关程度能够达到最大。

这两个要求表明，t1和u1应尽可能好的代表数据表X和Y,同时自变量的成分t1对因变量的成分u1又有最强的解释能力。

在第一个成分t1和u1被提取后，偏最小二乘回归分别实施X对t1的回归以及 Y对u1的回归。如果回归方程已经达到满意的精度，则算法终止；否则,将利用 X被t1解释后的残余信息以及Y 被t2解释后的残余信息进行第二轮的成分提取。如此往复，直到能达到一个较满意的精度为止。若最终对 X共提取了 m个成分t1,t2,...,tm，偏最小二乘回归将通过实施yk对t1,t2,...,tm的回归,然后再表达成yk关于原变量X1,X2,...,Xq的回归方程,k=1,2,...,p。

3.推导偏最小二乘回归：

为了彻底理解偏最小二乘回归，我建议下面的步骤你都亲自推导一遍。相信经过下面的推导，能让你对偏最小二乘

有一个更加清晰的认识。

- **step1:数据说明与标准化**

数据矩阵E0，F0，其中E0为自变量矩阵，每一行是一个样例，每一列代表了一个维度的变量；F0是因变量矩阵，解释同E0。

数据标准化即，要将数据中心化，方法是每个样本都做如下操作：减去一个维度变量的均值除以该维度的标准差。以下设E0，F0都为标准化了的数据。即：自变量经标准化处理后的数据矩阵记为E0（n*m），因变量经标准化处理后的数据矩阵记为F0（n*p）。

- **step2:求符合要求的主成分（☆）**

即求自变量与因变量的第一对主成分t1和u1，根据主成分原理，要求t1与u1的方差达到最大，这是因为：方差最大则表示的信息就越多。另一方面，又要求t1对u1有最大的解释能力，由典型相关分析的思路知，t1与u1的相关度达到最大值。

因此，综合上述两点，我们只要要求t1与u1的协方差达到最大，即：

Cov(t1,u1) –> max

而且，t1是X的线性组合，那么权重系数设为W1，即t1 = E0W1，同理，u1是Y的线性组合，u1 = F0C1。同时又要求，W1与C1同为单位向量，问题的数学表达式为：

max <E0w1,F0c1>

S.T.

||W1|| = 1;

||c1|| = 1

这就是一个条件极值的问题，你可以采用拉格朗日方法求解（如果你还有兴趣，可以查阅高数课本，当然，你也可以直接看结论，这里我只给出结论，推导省略，实际上推导并不影响你理解）。

通过拉格朗日求解，知w1就是矩阵E0'F0F0'E0的对应于最大特征值的特征向量，c1就是矩阵F0'E0E0'F0对应于最大特征值的最大特征向量，均单位化。

有了权系数w1，c1，自然可以求得主成分t1，u1。至此，第一对主成分完成。

- **step3:建立主成分与原自变量、因变量之间的回归（☆）**

建立E0，F0对t1，u1的三个回归方程，如下：

$$E_0 = t_1 p'_1 + E_1$$

$$F_0 = u_1 q'_1 + F^*_1$$

$$F_0 = t_1 r'_1 + F_1$$

式中，回归系数向量是：

$$p_1 = \frac{E_0' t_1}{\| t_1 \|^2}$$

$$q_1 = \frac{F_0' u_1}{\| u_1 \|^2}$$

$$r_1 = \frac{F_0' t_1}{\| t_1 \|^2}$$

而E1,F*1,F1分别是三个回归方程的残差矩阵.

- **step4:继续求主成分，直到满足要求**

用残差矩阵E1和F1取代E0和F0,然后,求第二个轴w2和c2以及第二个成分t2,u2,有

$$t_2 = E_1 w_2$$

$$u_2 = F_1 c_2$$

重新执行step3。直到求出所有主成分或者满足要求（后面说明）。

- **step5:推导因变量之于自变量的回归表达式（☆）**

如此经过step3-step4反复，若E0的秩为A，则可以求出：

$$E_0 = t_1 p_1' + \Lambda + t_A p_A'$$

$$F_0 = t_1 r_1' + \Lambda + t_A r_A' + F_A$$

由于t1....tA都可以表示E01，E02....E0q的线性组合，那么就自然还原成下面的形式：

$$y_k^* = \alpha_{k1} x_1^* + \Lambda + \alpha_{kp} x_p^* + F_{Ak} \qquad k=1,2,\cdots,q$$

Fak为残差矩阵Fa的第k列。这样，就求出了回归方程。

- **step6:检验-交叉有效性（☆）**

这是最后一步，也是非常重要的一步。下面要讨论的问题是在现有的数据表下,如何确定更好的回归方程。在许多情形下,偏最小二乘回归方程并不需要选用全部的成分进行回归建模,而是可以象在主成分分析一样,采用截尾的方式选择前m 个成分,仅用这m 个后续的成分就可以得到一个预测性较好的模型。事实上,如果后续的成分已经不能为解释因变量提供更有意义的信息时,采用过多的成分只会破坏对统计趋势的认识,引导错误的预测结论。

下面的问题是怎样来**确定所应提取的成分个数**。

在偏最小二乘回归建模中,究竟应该选取多少个成分为宜,这可通过考察增加一个新的成分后,能否对模型的

预测功能有明显的改进来考虑。采用类似于抽样测试法的工作方式,把所有n个样本点分成两部分:第一部分除去某个样本点i的所有样本点集合(共含n-1个样本点),用这部分样本点并使用h个成分拟合一个回归方程;第二部分是

$\hat{y}_{hj(-i)}$

把刚才被排除的样本点i代入前面拟合的回归方程,得到yj在样本点i上的拟合值 。对于每一个i=1,2,...,n,重复上述测试,则可以定义yj的预测误差平方和为**PRESShj**。有:

$$PRESS_{hj}=\sum_{i=1}^{n}(y_{ij}-\hat{y}_{hj(-i)})^2$$

定义Y 的预测误差平方和为**PRESSh**,有

$$PRESS_{h}=\sum_{j=1}^{p}PRESS_{hj}$$

显然,如果回归方程的稳健性不好,误差就很大,它对样本点的变动就会十分敏感,这种扰动误差的作用,就会加大PRESSh的值。

另外,再采用所有的样本点,拟合含h 个成分的回归方程。这是,记第i个样本点的预测值为$\hat{y}_{hji}$,则可以记yj的误差平方和为**SShj**,有

$$SS_{hj}=\sum_{i=1}^{n}(y_{ij}-\hat{y}_{hji})^2$$

定义Y的误差平方和为**SSh**,有

$$SS_{h}=\sum_{j=1}^{p}SS_{hj}$$

定义称为交叉有效性，对于每一个变量yk，定义

$$Q_{hk}^2=1-\frac{PRESS_{hk}}{SS_{(h-1)k}}$$

对于全部因变量Y,成分th**交叉有效性定义**为

$$Q_{h}^2=1-\frac{\sum_{k=1}^{q}PRESS_{hk}}{\sum SS_{(h-1)k}}=1-\frac{PRESS_{h}}{SS_{(h-1)}}$$

用交叉有效性测量成分th对预测模型精度的边际贡献有如下**两个尺度**。

$$Q_h^2\geq(1-0.95^2)=0.0975$$

(1) 当 时, th成分的边际贡献是显著的。显而易见, $Q_h^2 \geq 0.0975$ 与 $(PRESS_h / SS_{h-1}) < 0.95^2$ 是完全等价的决策原则。

(2) 对于k=1,2,...,q,至少有一个k,使得 $Q_h^2 \geq 0.0975$ 。这时增加成分th,至少使一个因变量yk的预测模型得到显著的改善,因此,也可以考虑增加成分th是明显有益的。

实现偏最小二乘回归算法步骤：

上面推导了偏最小二乘回归，分析了其中的原理。为了使得在实际应用中更加快速的使用偏最小二乘回归，在此，贴上实现偏最小二乘法实现的简洁步骤，需说明的是，下面算法来自司守奎老师《数学建模算法与应用》一书，该书推导过程跨度大，个人认为不适合新手直接阅读，建议你在理解了上述第二部分后再去阅读此书"偏最小二乘回归"章节，定会有更加高层次的认识。步骤如下：

（1）求矩阵 $E_0^T F_0 F_0^T E_0$ 最大特征值所对应的特征向量 w_1，求得成分得分向量 $\hat{t}_1 = E_0 w_1$，和残差矩阵 $E_1 = E_0 - \hat{t}_1 \alpha_1^T$，其中 $\alpha_1 = E_0^T \hat{t}_1 / \|\hat{t}_1\|^2$。

（2）求矩阵 $E_1^T F_0 F_0^T E_1$ 最大特征值所对应的特征向量 w_2，求得成分得分向量 $\hat{t}_2 = E_1 w_2$，和残差矩阵 $E_2 = E_1 - \hat{t}_2 \alpha_2^T$，其中 $\alpha_2 = E_1^T \hat{t}_2 / \|\hat{t}_2\|^2$。

⋮

（r）至第 r 步，求矩阵 $E_{r-1}^T F_0 F_0^T E_{r-1}$ 最大特征值所对应的特征向量 w_r，求得成分得分向量 $\hat{t}_r = E_{r-1} w_r$。

如果根据交叉有效性，确定共抽取 r 个成分 $t_1, \cdots, t_r$ 可以得到一个满意的预测模型，则求 F_0 在 $\hat{t}_1, \cdots, \hat{t}_r$ 上的普通最小二乘回归方程为

$$F_0 = \hat{t}_1 \beta_1^T + \cdots + \hat{t}_r \beta_r^T + F_r$$

把 $t_k = w_{k1}^* x_1 + \cdots + w_{km}^* x_m$（$k = 1,2,\cdots,r$），代入 $Y = t_1 \beta_1 + \cdots + t_r \beta_r$，即得 p 个因变量的偏最小二乘回归方程式

$$y_j = a_{j1} x_1 + \cdots + a_{jm} x_m, \quad (j = 1,2,\cdots,p)。$$

这里 $w_h^* = (w_{h1}^*, \cdots, w_{hm}^*)^T$ 满足 $\hat{t}_h = E_0 w_h^*$，$w_h^* = \prod_{j=1}^{h-1} (I - w_j \alpha_j^T) w_h$。

四、MATLAB实例以及实现

有必要贴出偏最小二乘的简单建模应用，并用matlab去是实现之，你可以按照上述步骤，通过基本的运算如求矩阵特征值等，来实现，也可以使用matlab工具箱方法实现之，下面给出的依旧是一个来自司守奎老师书本上的案例：

例： 采用兰纳胡德（Linnerud）给出的关于体能训练的数据进行偏小二乘回归建 模。在这个数据系统中被测的样本点，是某健身俱乐部的 20 位中年男子。被测变量分 为两组。第一组是身体特征指标 X，包括：体重、腰围、脉搏。第二组变量是训练结果指标Y，包括：单杠、弯曲、跳高。原始数据见表 1。 表 2 给出了这 6 个变量的简单相关系数矩阵。从相关系数矩阵可以看出，体重与 腰围是正相关的；体重、腰围与脉搏负相关；而在单杠、弯曲与跳高之间是正相关的。 从两组变量间的关系看，单杠、弯曲和跳高的训练成绩与体重、腰围负相关，与脉搏正相关。

表 1　体能训练数据

No	体重(x_1)	腰围(x_2)	脉搏(x_3)	单杠(y_1)	弯曲(y_2)	跳高(y_3)
1	191	36	50	5	162	60
2	189	37	52	2	110	60
3	193	38	58	12	101	101
4	162	35	62	12	105	37
5	189	35	46	13	155	58
6	182	36	56	4	101	42
7	211	38	56	8	101	38
8	167	34	60	6	125	40
9	176	31	74	15	200	40
10	154	33	56	17	251	250
11	169	34	50	17	120	38
12	166	33	52	13	210	115
13	154	34	64	14	215	105
14	247	46	50	1	50	50
15	193	36	46	6	70	31
16	202	37	62	12	210	120
17	176	37	54	4	60	25
18	157	32	52	11	230	80
19	156	33	54	15	225	73
20	138	33	68	2	110	43
均值	178.6	35.4	56.1	9.45	145.55	70.3
标准差	24.6905	3.202	7.2104	5.2863	62.5666	51.2775

表 2　相关系数矩阵

1	0.8702	−0.3658	−0.3897	−0.4931	−0.2263
0.8702	1	−0.3529	−0.5522	−0.6456	−0.1915
−0.3658	−0.3529	1	0.1506	0.225	0.0349
−0.3897	−0.5522	0.1506	1	0.6957	0.4958
−0.4931	−0.6456	0.225	0.6957	1	0.6692
−0.2263	−0.1915	0.0349	0.4958	0.6692	1

可以利用如下的MATLAB程序：

```
clc,clear
load pz.txt %原始数据存放在纯文本文件 pz.txt 中
mu=mean(pz);sig=std(pz); %求均值和标准差
rr=corrcoef(pz); %求相关系数矩阵
data=zscore(pz); %数据标准化,变量记做 X*和 Y*
n=3;m=3; %n 是自变量的个数,m 是因变量的个数
x0=pz(:,1:n);y0=pz(:,n+1:end); %原始的自变量和因变量数据
e0=data(:,1:n);f0=data(:,n+1:end); %标准化后的自变量和因变量数据
-679-
num=size(e0,1);%求样本点的个数
chg=eye(n); %w 到 w*变换矩阵的初始化
for i=1:n
%以下计算 w，w*和 t 的得分向量，
matrix=e0'*f0*f0'*e0;
[vec,val]=eig(matrix); %求特征值和特征向量
val=diag(val); %提出对角线元素，即提出特征值
[val,ind]=sort(val,'descend');
w(:,i)=vec(:,ind(1)); %提出最大特征值对应的特征向量
w_star(:,i)=chg*w(:,i); %计算 w*的取值
t(:,i)=e0*w(:,i); %计算成分 ti 的得分
alpha=e0'*t(:,i)/(t(:,i)'*t(:,i)); %计算 alpha_i
chg=chg*(eye(n)-w(:,i)*alpha'); %计算 w 到 w*的变换矩阵
e=e0-t(:,i)*alpha'; %计算残差矩阵
e0=e;
%以下计算 ss(i)的值
beta=t\f0; %求回归方程的系数，数据标准化，没有常数项
cancha=f0-t*beta; %求残差矩阵
```

```
ss(i)=sum(sum(cancha.^2)); %求误差平方和
%以下计算 press(i)
for j=1:num
t1=t(:,1:i);f1=f0;
she_t=t1(j,:);she_f=f1(j,:); %把舍去的第 j 个样本点保存起来
t1(j,:)=[];f1(j,:)=[]; %删除第 j 个观测值
beta1=[t1,ones(num-1,1)]\f1; %求回归分析的系数,这里带有常数项
cancha=she_f-she_t*beta1(1:end-1,:)-beta1(end,:); %求残差向量
press_i(j)=sum(cancha.^2); %求误差平方和
end
press(i)=sum(press_i);
Q_h2(1)=1;
if i>1, Q_h2(i)=1-press(i)/ss(i-1); end
if Q_h2(i)<0.0975
fprintf('提出的成分个数 r=%d',i); break
end
end
beta_z=t\f0; %求 Y*关于 t 的回归系数
xishu=w_star*beta_z; %求 Y*关于 X*的回归系数，每一列是一个回归方程
mu_x=mu(1:n);mu_y=mu(n+1:end); %提出自变量和因变量的均值
sig_x=sig(1:n);sig_y=sig(n+1:end); %提出自变量和因变量的标准差
ch0=mu_y-(mu_x./sig_x*xishu).*sig_y; %计算原始数据回归方程的常数项
for i=1:m
xish(:,i)=xishu(:,i)./sig_x'*sig_y(i); %计算原始数据回归方程的系数
end
sol=[ch0;xish] %显示回归方程的系数，每一列是一个方程，每一列的第一个数是常数项
save mydata x0 y0 num xishu ch0 xish
```

求解过程如下：

计算得只要提出两个成分 t_1,t_2 即可，交叉有效性 $Q_2^2=-0.3263$。w_h 与 w_h^* 的取值见表 3，成分 t_h 的得分 $\hat{t}_h$ 见表 4。

表 3　w_h 与 w_h^* 的取值

自变量	w_1	w_2	w_1^*	w_2^*
x_1	−0. 5899	-0.4688	−0. 5899	-0.3679
x_2	−0. 7713	0.5680	−0. 7713	0.6999
x_3	0. 2389	0.6765	0. 2389	0.6356

表 4　成分 t_h 的得分 $\hat{t}_h$

No	1	2	3	4	5	6	7	8	9	10
$\hat{t}_1$	−0. 6429	−0. 7697	−0. 9074	0. 6884	−0. 4867	−0. 2291	−1. 4037	0. 7436	1. 7151	1. 1626
$\hat{t}_2$	−0. 5914	−0. 1667	0. 5212	0. 68	−1. 1328	0. 0717	0. 0767	0. 2106	0. 6549	−0. 1668
No	11	12	13	14	15	16	17	18	19	20
$\hat{t}_1$	0. 3645	0. 7433	1. 1867	−4. 3898	−0. 8232	−0. 749	−0. 3929	1. 1993	1. 0485	1. 9424
$\hat{t}_2$	−0. 7007	−0. 6983	0. 757	0. 76	−0. 9738	0. 5211	0. 2034	−0. 7827	−0. 3729	1. 1294

标准化变量 $\tilde{y}_k$ 关于成分 t_1,t_2 的回归模型如下

$$\tilde{y}_k=r_{1k}t_1+r_{2k}t_2,\quad k=1,2,3$$

由于成分 t_h 可以写成原变量的标准化变量 $\tilde{x}_j$ 的函数，即有

$$t_h=w_{1h}^*\tilde{x}_1+w_{2h}^*\tilde{x}_2+w_{3h}^*\tilde{x}_3,\quad h=1,2$$

由此可得由成分 t_1,t_2 所建立的偏最小二乘回归模型为

$$\begin{aligned}\tilde{y}_k&=r_{1k}(w_{11}^*\tilde{x}_1+w_{21}^*\tilde{x}_2+w_{31}^*\tilde{x}_3)+r_{2k}(w_{12}^*\tilde{x}_1+w_{22}^*\tilde{x}_2+w_{32}^*\tilde{x}_3)\\&=(r_{1k}w_{11}^*+r_{2k}w_{12}^*)\tilde{x}_1+(r_{1k}w_{21}^*+r_{2k}w_{22}^*)\tilde{x}_2+(r_{1k}w_{31}^*+r_{2k}w_{32}^*)\tilde{x}_3\end{aligned}$$

有关 $r_h=(r_{h1},r_{h2},r_{h3})$ 的计算结果见表 5。

表 5　回归系数 r_h

k	1	2	3
r_1	0.3416	0.4161	0.1430
r_2	−0. 3364	−0. 2908	−0. 0652

所以，有

$$\tilde{y}_1 = -0.0778\tilde{x}_1 - 0.4989\tilde{x}_2 - 0.1322\tilde{x}_3$$
$$\tilde{y}_2 = -0.1385\tilde{x}_1 - 0.5244\tilde{x}_2 - 0.0854\tilde{x}_3$$
$$\tilde{y}_3 = -0.0604\tilde{x}_1 - 0.1559\tilde{x}_2 - 0.0073\tilde{x}_3$$

将标准化变量 $\tilde{y}_k, \tilde{x}_k (k = 1,2,3)$ 分别还原成原始变量 $y_k, x_k (k = 1,2,3)$，则回归方程为

$$y_1 = 47.0197 - 0.0167x_1 - 0.8237x_2 - 0.0969x_3$$
$$y_2 = 612.5671 - 0.3509x_1 - 10.2477x_2 - 0.7412x_3$$
$$y_3 = 183.9849 - 0.1253x_1 - 2.4969x_2 - 0.0518x_3$$

为了更直观、迅速地观察各个自变量在解释 $y_k (k = 1,2,3)$ 时的边际作用，可以绘制回归系数图，见图 1。这个图是针对标准化数据的回归方程的。

从回归系数图中可以立刻观察到，腰围变量在解释三个回归方程时起到了极为重要的作用。然而，与单杠及弯曲相比，跳高成绩的回归方程显然不够理想，三个自变量对它的解释能力均很低。

为了考察这三个回归方程的模型精度，我们以 $(\hat{y}_{ik}, y_{ik})$ 为坐标值，对所有的样本点绘制预测图。$\hat{y}_{ik}$ 是第 k 个变量，第 i 个样本点 (y_{ik}) 的预测值。在这个预测图上，如果所有点都能在图的对角线附近均匀分布，则方程的拟合值与原值差异很小，这个方程的拟合效果就是满意的。体能训练的预测图见图 2。

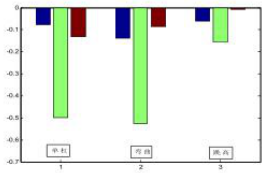

图 1　回归系数的直方图

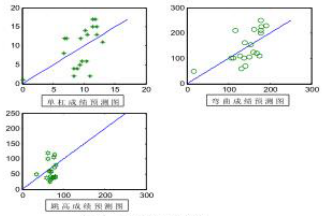

图 2　体能训练预测图

至此，偏最小二乘回归推导以及案例讲解完毕，通过第二部分可以了解片最小二乘回归的原理，这是论文正确写作的保证，参考第三部分可以使得你在具体应用中快速实现之，这是正确求解的保证。

参考：

《数学建模算法与引用》司守奎老师

pdf下载链接：

http://vdisk.weibo.com/s/t0L2pU6fgiP9L?category_id=0&parents_ref=t0L2pU6fgiPav

以上，请批评指正。

分类: 算法

标签: 数学建模, matlab

Y.D
关注 - 3
粉丝 - 30

« 上一篇： 【建模应用】PCA主成分分析原理详解
» 下一篇： 【Oh】论宇宙是怎么来的

posted @ 2018-05-13 11:13 Y.D 阅读(24288) 评论(0) 编辑 收藏

积分 - 61621
排名 - 10449

阅读排行榜

1. 【Python】一份非常好的Matplotlib教程(37361)
2. 【插值】插值方法原理详解(26618)
3. 【建模应用】PLS偏最小二乘回归原理与应用(24288)
4. 【pytorch】pytorch-LSTM(5677)
5. 【C/C++】动态内存分配和链表(3149)