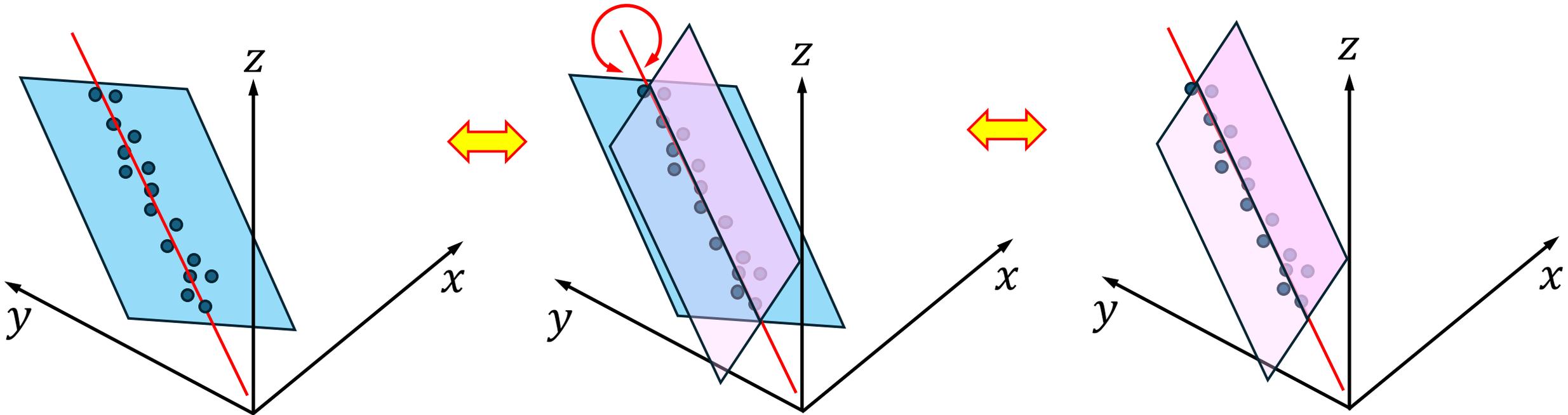


PLS回帰分析

- ・主成分と目的変数との間の共分散が最大になるように主成分を抽出
- ・抽出した主成分をもとに線形回帰実施
- ・多重共線性(2つ以上の説明変数が高い相関)に対応可能
 - ①高い相関がある列を削除
 - ②共線性がある列をまとめ、PLS回帰分析

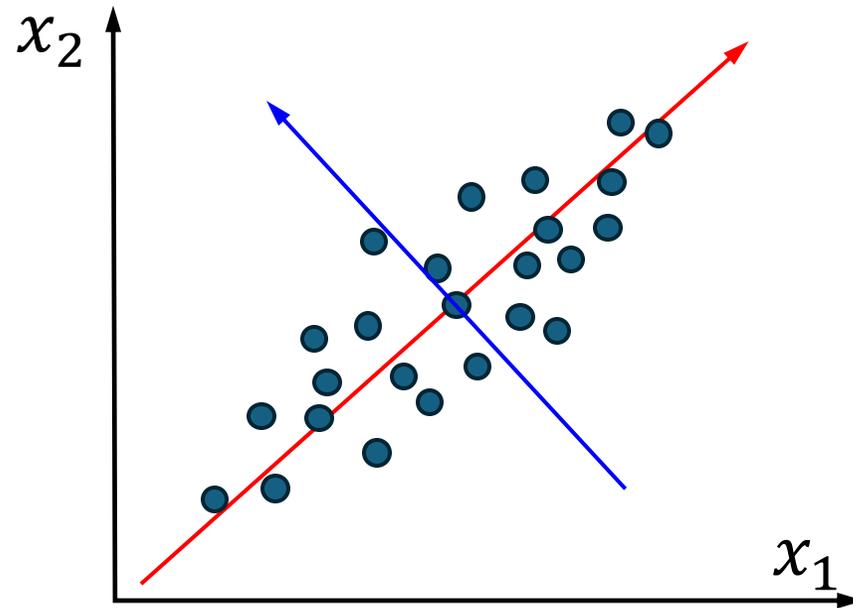
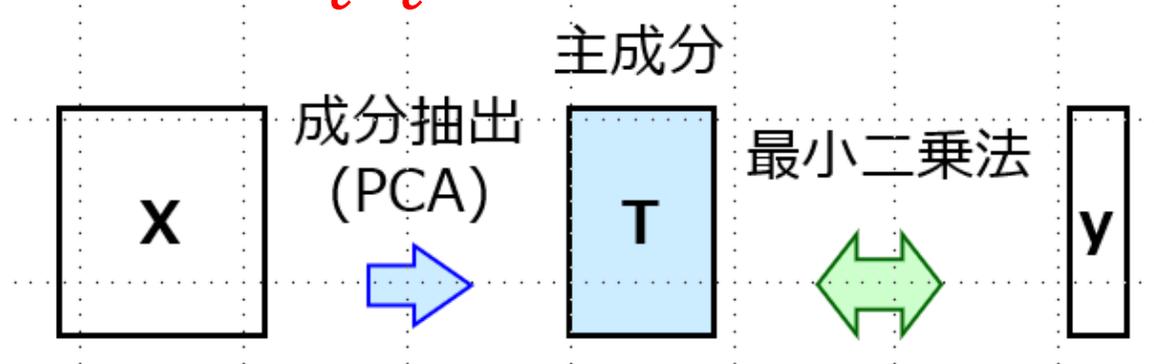
高い相関があると、その軸の周りに、無数の平面の可能性が生じ、一意的に定まらない



主成分分析 (PCA)

射影した点の**分散**が最大になる主成分 t を作る

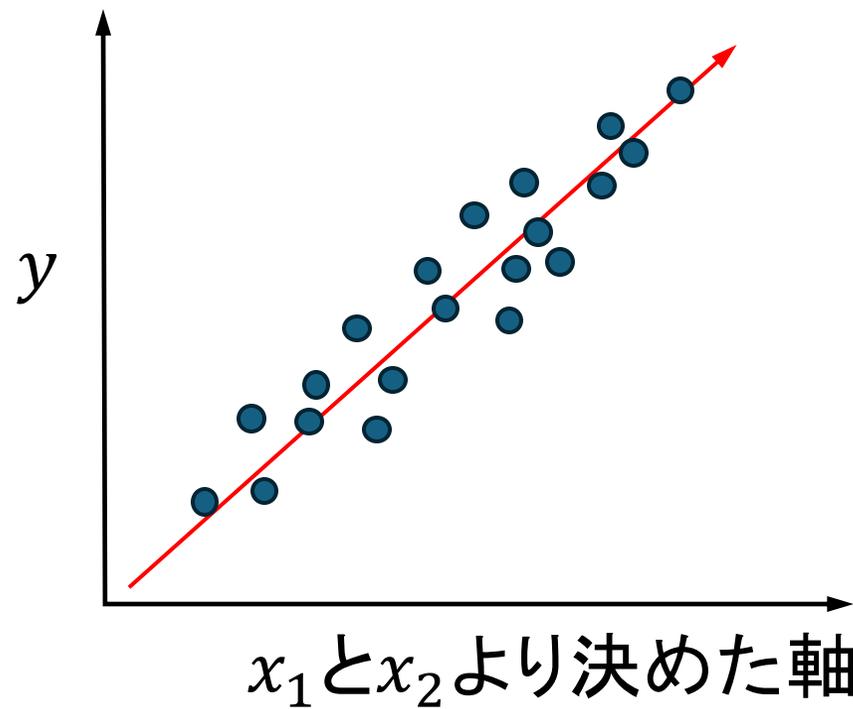
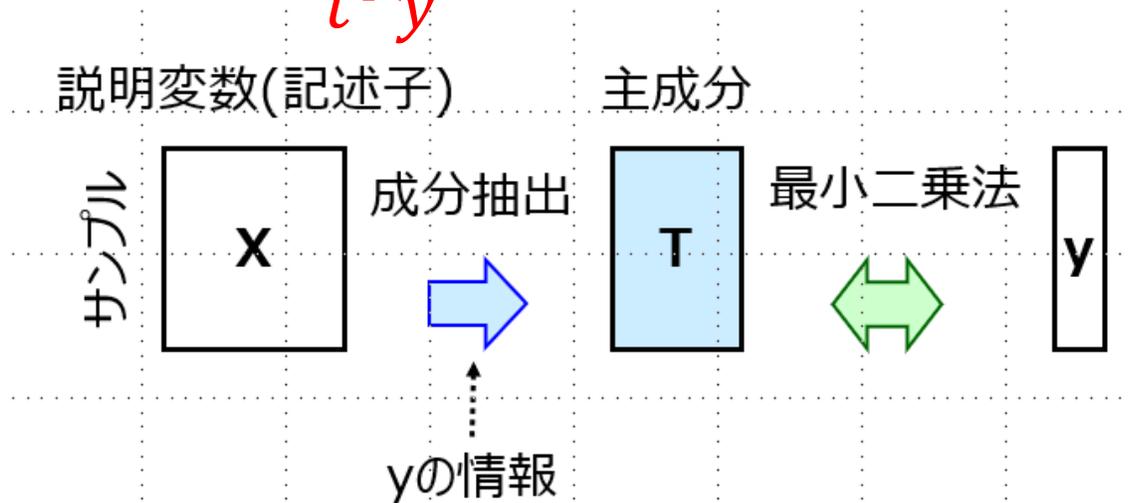
$$t^T t$$



部分的最小二乗回帰 (PSL)

目的変数 y との**分散**が最大になる主成分 t を作る

$$t^T y$$



Minitab

- 60個の大豆粉のサンプルを分析
- 各サンプルの水分と脂肪分を調べ、88波長で近赤外線(NIR)スペクトルデータを記録
- 60サンプルからランダムに54サンプルを選択し、PLS回帰を使用して応答変数(水分と脂肪分)と予測変数(88 NIR波長)との関係を推定
- 残りの残り6つのサンプルは、モデルの予測能力を評価するために検定のデータセットとして使用

88波長データ

水分データ

脂肪データ

予測用波長データ

54サンプル

水分データ

脂肪データ

C1	C2	C3	C4	C5	C6	C7	C8	C9	C10	C11	
1	-0.0234600	-0.0223210	-0.0113710	0.0176180	0.051994	0.073911	0.089845	0.088451	0.060333	0.039534	0.0298130
2	0.0041220	0.0058260	0.0191230	0.0542480	0.093939	0.117926	0.135243	0.133215	0.101889	0.079112	0.0683850
3	-0.0005440	0.0012330	0.0147100	0.0500820	0.090107	0.114502	0.131736	0.129284	0.097622	0.074715	0.0639950
4	0.0158820	0.0178300	0.0316620	0.0680940	0.109570	0.135648	0.152658	0.148927	0.117673	0.095343	0.0847330
5	-0.0142650	-0.0134880	-0.0033880	0.0243600	0.057469	0.078266	0.093646	0.092370	0.065371	0.045312	0.0358140
6	-0.0216550	-0.0202930	-0.0088610	0.0213920	0.056366	0.077960	0.093283	0.091218	0.063723	0.043710	0.0342380
7	-0.0048730	-0.0037370	0.0081070	0.0399640	0.077623	0.100209	0.117512	0.116463	0.086306	0.063829	0.0531770
8	-0.0237200	-0.0230960	-0.0133890	0.0137260	0.046695	0.066931	0.082532	0.081704	0.054663	0.034524	0.0250080
9	0.0074040	0.0092800	0.0232680	0.0600970	0.102761	0.128518	0.147556	0.145833	0.112283	0.087851	0.0764770
10	0.0170530	0.0192130	0.0342860	0.0733930	0.117222	0.144561	0.162751	0.159238	0.126221	0.102554	0.0913220
11	-0.0237850	-0.0224660	-0.0106580	0.0191280	0.053247	0.075975	0.090426	0.087255	0.061199	0.042285	0.0333080
12	0.0146560	0.0162710	0.0286960	0.0618190	0.098987	0.122909	0.137947	0.133985	0.105607	0.085620	0.0760720
13	-0.0288570	-0.0279170	-0.0176260	0.0098450	0.042736	0.063914	0.079507	0.078715	0.052486	0.032807	0.0234680
14	0.0084400	0.0103030	0.0240130	0.0606820	0.103392	0.129177	0.148476	0.146921	0.113116	0.088463	0.0769790
15	-0.0441970	-0.0434820	-0.0340800	-0.0110110	0.015761	0.034351	0.045969	0.043146	0.021929	0.006477	-0.0008490
16	0.0098790	0.0110360	0.0224110	0.0536500	0.089896	0.112724	0.128557	0.126118	0.097183	0.076372	0.0665290
17	-0.0341740	-0.0335230	-0.0241240	-0.0004190	0.027194	0.045955	0.057758	0.054837	0.032940	0.017091	0.0095720
18	-0.0001790	0.0017380	0.0150130	0.0504060	0.092174	0.117708	0.136766	0.135027	0.101265	0.076819	0.0654980
19	-0.0324080	-0.0315570	-0.0215970	0.0035990	0.033097	0.052688	0.065639	0.063280	0.039855	0.022669	0.0145690
20	0.0041720	0.0060240	0.0194470	0.0541020	0.093107	0.117556	0.134088	0.131456	0.101350	0.079606	0.0694080

C84	C85	C86	C87	C88	C89	C90	C91	C92	C93	C94
84	85	86	87	88	水分	脂肪	テスト1	テスト2	テスト3	テスト4
0.652913	0.671835	0.69610	0.71066	0.71394	9.9000	22.4000	-0.0025500	-0.0004820	0.0134620	0.0496850
0.814006	0.836222	0.86377	0.88062	0.88615	14.0000	19.8000	-0.0206480	-0.0196680	-0.0089540	0.0181610
0.795355	0.816519	0.84318	0.85897	0.86398	13.7000	19.7000	-0.0027640	-0.0012940	0.0106370	0.0431200
0.895322	0.919610	0.94950	0.96801	0.97490	15.0000	15.8000	0.0490530	0.0499600	0.0600980	0.0885710
0.664643	0.684335	0.70908	0.72437	0.72835	11.4000	22.2000	0.0094850	0.0116090	0.0262940	0.0643190
0.679385	0.699631	0.72504	0.74080	0.74528	10.9000	17.1000	-0.0267590	-0.0256650	-0.0147530	0.0137050
0.724381	0.745141	0.77109	0.78779	0.79283	9.0000	20.7000				
0.637801	0.657243	0.68195	0.69754	0.70156	8.3000	20.5000				
0.857380	0.879703	0.90724	0.92467	0.92986	12.9000	19.7000				
0.911601	0.935618	0.96443	0.98243	0.98886	15.4000	15.9000				
0.652017	0.671374	0.69569	0.71002	0.71370	10.2000	14.7000				
0.863879	0.889085	0.91991	0.93822	0.94501	17.2000	15.6000				
0.642886	0.663447	0.68896	0.70496	0.70889	9.5000	22.3000				
0.874177	0.897714	0.92673	0.94569	0.95165	13.3000	21.0000				
0.482807	0.498204	0.51898	0.53021	0.53185	6.0000	16.6000				
0.804365	0.827630	0.85601	0.87356	0.87977	15.5000	19.7000				
0.535927	0.553812	0.57688	0.58983	0.59219	9.0000	16.4000				
0.820777	0.840835	0.86631	0.88109	0.88484	12.9000	21.4000				
0.555856	0.573730	0.59688	0.61029	0.61322	6.3000	17.8000				
0.782121	0.803328	0.82935	0.84544	0.85038	13.3000	19.1000				

C176	C177	C178	C179	C180	C181
テスト86	テスト87	テスト88	水分2	脂肪2	
0.850987	0.867328	0.872910	13.5000	18.1000	
0.668220	0.682408	0.685809	9.0000	15.6000	
0.869857	0.887175	0.892366	14.9000	19.8000	
0.918884	0.936982	0.943686	18.4000	16.3000	
0.914459	0.931544	0.937531	15.8000	16.2000	
0.659287	0.672916	0.675679	9.9000	22.5000	

水分を目的変数、88波長のデータを説明変数として重回帰分析をした結果

次の項は推定できないため、取り除かれました:

54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88

回帰式

水分 = 31.74 + 12557 1 + 33719 2 - 105433 3 + 104836 4 - 101133 5 + 71405 6 + 20705 7 - 20723 8
+ 79116 9 - 362445 10 + 480306 11 - 213681 12 - 77013 13 + 29951 14 + 153483 15
- 171572 16 + 73007 17 + 45284 18 - 52467 19 + 251.4 20 - 31614 21 + 95064 22 - 54991 23
- 33657 24 + 32018 25 - 1458 26 - 8203 27 + 31671 28 + 21145 29 - 107399 30 + 130296 31
- 120673 32 + 112733 33 - 46275 34 - 130661 35 + 150038 36 - 29544 37 - 7184 38
- 29302 39 + 10595 40 - 31253 41 + 62225 42 - 28927 43 + 2399 44 + 76282 45 - 31015 46
- 72545 47 + 10142 48 + 44275 49 - 21479 50 + 27575 51 - 45886 52 + 25741 53

係数

項	係数	係数の標準誤差	t値	p値	VIF
定数	31.74	*	*	*	
1	12557	*	*	*	14027959.79
2	33719	*	*	*	2.20888E+08
3	-105433	*	*	*	1.72170E+08
4	104836	*	*	*	9.41736E+08
5	-101133	*	*	*	1.89874E+09
6	71405	*	*	*	1.03990E+08
7	20705	*	*	*	2.43238E+08
8	-20723	*	*	*	2.00814E+08
9	79116	*	*	*	1.05923E+09
10	-362445	*	*	*	1.80703E+09
11	480306	*	*	*	4.56618E+09
12	-213681	*	*	*	1.35843E+10
13	-77013	*	*	*	1.08944E+10
14	29951	*	*	*	5.83901E+09
15	153483	*	*	*	1.85040E+09
16	-171572	*	*	*	2.99032E+08

重回帰分析不能

- データ全て使用できない
- 係数のp値が算出不能
- VIF (Variance Inflation Factor) が10以上のとき
多重共線性あり

①標本データ「大豆粉.MTW」を開く

ファイル(F) 編集(E) データ(A) 計算(C) 統計(S) グラフ(G) ビュー(V) ヘルプ(H) アシスタント(N) 予測分析モジュール(P) その他のツール(T)

	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10	C11
	1	2	3	4	5	6	7	8	9	10	11
1	-0.0234600	-0.0223210	-0.0113710	0.0176180	0.051994	0.073911	0.089845	0.088451	0.060333	0.039534	0.0298130
2	0.0041220	0.0058260	0.0191230	0.0542480	0.093939	0.117926	0.135243	0.133215	0.101889	0.079112	0.0683850
3	-0.0005440	0.0012330	0.0147100	0.0500820	0.090107	0.114502	0.131736	0.129284	0.097622	0.074715	0.0639950
4	0.0158820	0.0178300	0.0316620	0.0680940	0.109570	0.135648	0.152658	0.148927	0.117673	0.095343	0.0847330
5	-0.0142650	-0.0134880	-0.0033880	0.0243600	0.057469	0.078266	0.093646	0.092370	0.065371	0.045312	0.0358140
6	-0.0216550	-0.0202930	-0.0088610	0.0213920	0.056366	0.077960	0.093283	0.091218	0.063723	0.043710	0.0342380
7	-0.0048730	-0.0037370	0.0081070	0.0399640	0.077623	0.100209	0.117512	0.116463	0.086306	0.063829	0.0531770
8	-0.0237200	-0.0230960	-0.0133890	0.0137260	0.046695	0.066931	0.082532	0.081704	0.054663	0.034524	0.0250080
9	0.0074040	0.0092800	0.0232680	0.0600970	0.102761	0.128518	0.147556	0.145833	0.112283	0.087851	0.0764770
10	0.0170530	0.0192130	0.0342860	0.0733930	0.117222	0.144561	0.162751	0.159238	0.126221	0.102554	0.0913220
11	-0.0237850	-0.0224660	-0.0106580	0.0191280	0.053247	0.075975	0.090426	0.087255	0.061199	0.042285	0.0333080
12	0.0146560	0.0162710	0.0286960	0.0618190	0.098987	0.122909	0.137947	0.133985	0.105607	0.085620	0.0760720
13	-0.0288570	-0.0279170	-0.0176260	0.0098450	0.042736	0.063914	0.079507	0.078715	0.052486	0.032807	0.0234680
14	0.0084400	0.0103030	0.0240130	0.0606820	0.103392	0.129177	0.148476	0.146921	0.113116	0.088463	0.0769790
15	-0.0441970	-0.0434820	-0.0340800	-0.0110110	0.015761	0.034351	0.045969	0.043146	0.021929	0.006477	-0.0008490
16	0.0098790	0.0110360	0.0224110	0.0536500	0.089896	0.112724	0.128557	0.126118	0.097183	0.076372	0.0665290
17	-0.0341740	-0.0335230	-0.0241240	-0.0004190	0.027194	0.045955	0.057758	0.054837	0.032940	0.017091	0.0095720
18	-0.0001790	0.0017380	0.0150130	0.0504060	0.092174	0.117708	0.136766	0.135027	0.101265	0.076819	0.0654980
19	-0.0324080	-0.0315570	-0.0215970	0.0035990	0.033097	0.052688	0.065639	0.063280	0.039855	0.022669	0.0145690
20	0.0041720	0.0060240	0.0194470	0.0541020	0.093107	0.117556	0.134088	0.131456	0.101350	0.079606	0.0694080

ファイル(F) 編集(E) データ(A) 計算(C) 統計(S) グラフ(G) ビュー(V) ヘルプ(H) アシスタント(N) 予測分析モジュール(P) その他のツール(T)

C84	C85	C86	C87	C88	C89	C90	C91	C92	C93	C94
84	85	86	87	88	水分	脂肪	テスト1	テスト2	テスト3	テスト4
0.652913	0.671835	0.69610	0.71066	0.71394	9.9000	22.4000	-0.0025500	-0.0004820	0.0134620	0.0496850
0.814006	0.836222	0.86377	0.88062	0.88615	14.0000	19.8000	-0.0206480	-0.0196680	-0.0089540	0.0181610
0.795355	0.816519	0.84318	0.85897	0.86398	13.7000	19.7000	-0.0027640	-0.0012940	0.0106370	0.0431200
0.895322	0.919610	0.94950	0.96801	0.97490	15.0000	15.8000	0.0490530	0.0499600	0.0600980	0.0885710
0.664643	0.684335	0.70908	0.72437	0.72835	11.4000	22.2000	0.0094850	0.0116090	0.0262940	0.0643190
0.679385	0.699631	0.72504	0.74080	0.74528	10.9000	17.1000	-0.0267590	-0.0256650	-0.0147530	0.0137050
0.724381	0.745141	0.77109	0.78779	0.79283	9.0000	20.7000				
0.637801	0.657243	0.68195	0.69754	0.70156	8.3000	20.5000				
0.857380	0.879703	0.90724	0.92467	0.92986	12.9000	19.7000				
0.911601	0.935618	0.96443	0.98243	0.98886	15.4000	15.9000				
0.652017	0.671374	0.69569	0.71002	0.71370	10.2000	14.7000				
0.863879	0.889085	0.91991	0.93822	0.94501	17.2000	15.6000				
0.642886	0.663447	0.68896	0.70496	0.70889	9.5000	22.3000				
0.874177	0.897714	0.92673	0.94569	0.95165	13.3000	21.0000				
0.482807	0.498204	0.51898	0.53021	0.53185	6.0000	16.6000				
0.804365	0.827630	0.85601	0.87356	0.87977	15.5000	19.7000				
0.535927	0.553812	0.57688	0.58983	0.59219	9.0000	16.4000				
0.820777	0.840835	0.86631	0.88109	0.88484	12.9000	21.4000				
0.555856	0.573730	0.59688	0.61029	0.61322	6.3000	17.8000				
0.782121	0.803328	0.82935	0.84544	0.85038	13.3000	19.1000				

ファイル(F) 編集(E) データ(A) 計算(C) 統計(S) グラフ(G) ビュー(V) ヘルプ(H) アシスタント(N) 予測分析モジュール(P) その他のツール(T)

C176	C177	C178	C179	C180	C181
テスト86	テスト87	テスト88	水分2	脂肪2	
0.850987	0.867328	0.872910	13.5000	18.1000	
0.668220	0.682408	0.685809	9.0000	15.6000	
0.869857	0.887175	0.892366	14.9000	19.8000	
0.918884	0.936982	0.943686	18.4000	16.3000	
0.914459	0.931544	0.937531	15.8000	16.2000	
0.659287	0.672916	0.675679	9.9000	22.5000	



②統計 → 回帰 → PLS をクリック

ファイル(F) 編集(E) データ(A) 計算(C) 統計(S) グラフ(G) ビュー(V) ヘルプ(H) アシスタント(N) 予測分析モジュール

- 基本統計(B)
- 回帰(R)
 - 適合線プロット(F)...
 - 回帰(R)
 - 非線形回帰(E)...
 - 安定性分析(S)
 - 直交回帰(T)...
 - PLS(P)...**
 - 2値適合線プロット(B)...
- 分散分析(A)
- 実験計画法 (DOE) (D)
- 管理図(C)
- 品質ツール(Q)
- 信頼性/生存時間(L)
- 予測分析(V)
- 多変量解析(M)

③応答に、水分・脂肪を選択 モデルに'1'-'88' 予測をクリック

PLS

応答(E): 水分, 脂肪

モデル(D): '1'-'88'

カテゴリ予測変数(オプション)(C):

最大成分数(M):

オプション(T)... 予測(P)...

選択 グラフ(G)... 結果(R)... 保存(S)...

ヘルプ OK(O) キャンセル

④連続予測変数の新しい観測値に、 テスト1-テスト88を選択 応答の新しい観測値(オプション)に、 水分2脂肪2を選択 OKをクリック

PLS: 予測

C166 テスト76
C167 テスト77
C168 テスト78
C169 テスト79
C170 テスト80
C171 テスト81
C172 テスト82
C173 テスト83
C174 テスト84
C175 テスト85
C176 テスト86
C177 テスト87
C178 テスト88
C179 水分2
C180 脂肪2

連続予測変数の新しい観測値(N): 'テスト1'-'テスト88'

カテゴリ予測変数の新しい観測値(E):

応答の新しい観測値(オプション)(R): '水分2' '脂肪2'

信頼水準(L): 95

保存
 適合値(F)
 適合値の標準誤差(S)
 信頼限界(C)
 予測限界(P)

選択

ヘルプ OK(O) キャンセル

⑤「グラフ」をクリック

⑦ビュー → 出力のみ



⑥見たいグラフにチェック入れてOK

方法

交差検証 なし
 計算する成分 設定
 計算された成分数 10

実行結果

水分の分散分析

要因	自由度	平方和	平均平方	F値	p値
回帰	10	468.516	46.8516	61.46	0.000
残差誤差	43	32.777	0.7623		
合計	53	501.293			

脂肪の分散分析

要因	自由度	平方和	平均平方	F値	p値
回帰	10	266.378	26.6378	36.89	0.000
残差誤差	43	31.050	0.7221		
合計	53	297.428			

水分に対するモデル選択および検証

成分	X分散	誤差	R二乗
1	0.984976	96.9288	0.806643
2	0.996400	88.9900	0.822479
3	0.997757	71.9304	0.856510
4	0.999427	58.3174	0.883666
5	0.999722	58.1261	0.884048
6	0.999853	48.5236	0.903203
7	0.999963	45.9824	0.908272
8	0.999976	33.1545	0.933862
9	0.999982	32.8074	0.934554
10	0.999986	32.7773	0.934615

脂肪に対するモデル選択および検証

成分	X分散	誤差	R二乗
1	0.984976	282.519	0.050127
2	0.996400	229.964	0.226824
3	0.997757	115.951	0.610155
4	0.999427	98.285	0.669550
5	0.999722	57.994	0.805015
6	0.999853	53.097	0.821480
7	0.999963	52.010	0.825133
8	0.999976	48.842	0.835784
9	0.999982	34.344	0.884529
10	0.999986	31.050	0.895604

水分に対するモデルを使用する新しい観測値に対する予測応答

行	適合値	適合値の標準誤差	95%信頼区間	95%予測区間
1	14.5184	0.388841	(13.7343, 15.3026)	(12.5910, 16.4459)
2	9.3049	0.372712	(8.5532, 10.0565)	(7.3904, 11.2193)
3	14.1790	0.504606	(13.1614, 15.1966)	(12.1454, 16.2127)
4	16.4477	0.559704	(15.3189, 17.5764)	(14.3562, 18.5391)
5	15.1872	0.358044	(14.4652, 15.9093)	(13.2842, 17.0903)
6	9.4639	0.485613	(8.4846, 10.4433)	(7.4492, 11.4787)

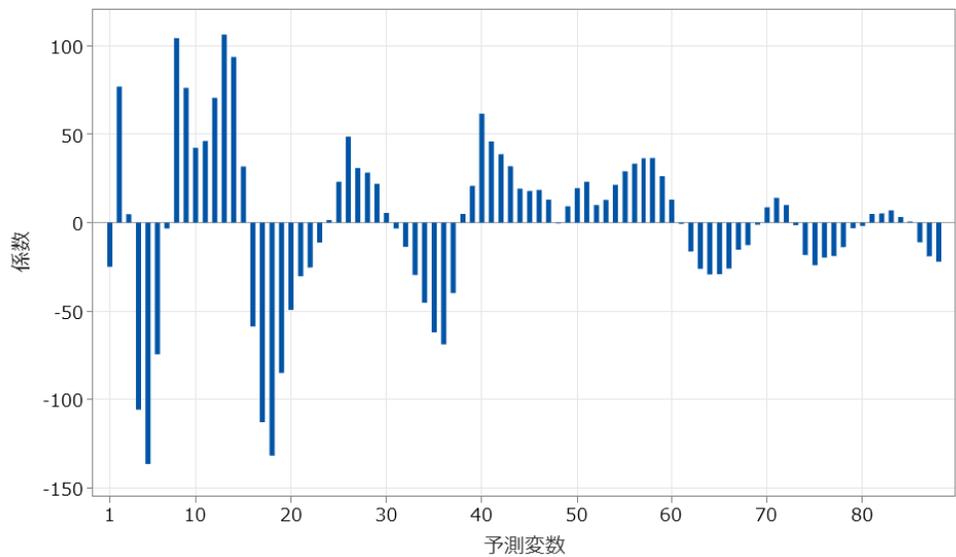
テストR二乗: 0.906451

脂肪に対するモデルを使用する新しい観測値に対する予測応答

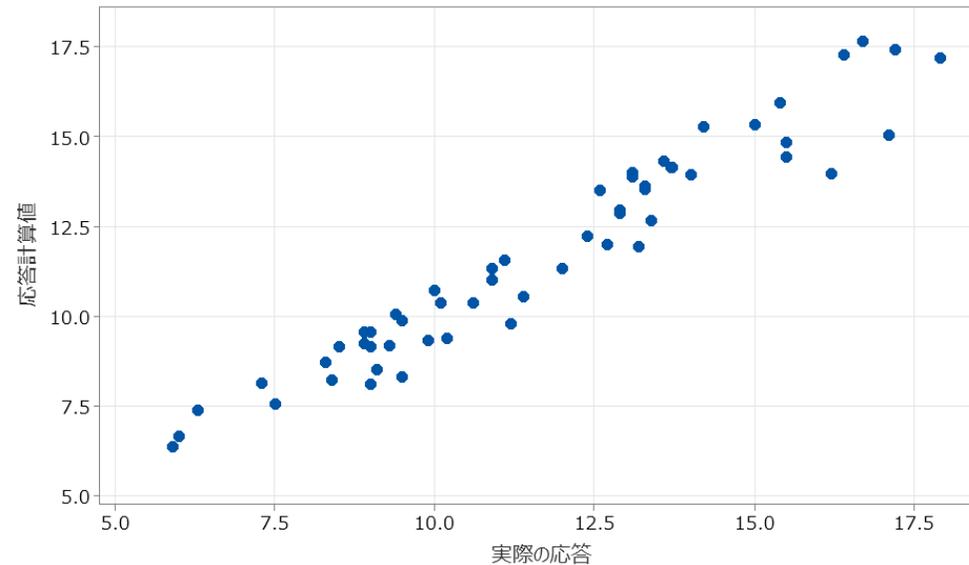
行	適合値	適合値の標準誤差	95%信頼区間	95%予測区間
1	18.7372	0.378459	(17.9740, 19.5004)	(16.8612, 20.6132)
2	15.3782	0.362762	(14.6466, 16.1098)	(13.5149, 17.2415)
3	20.7838	0.491134	(19.7933, 21.7743)	(18.8044, 22.7632)
4	14.3684	0.544761	(13.2698, 15.4670)	(12.3328, 16.4040)
5	16.6016	0.348485	(15.8988, 17.3044)	(14.7494, 18.4538)
6	20.7471	0.472648	(19.7939, 21.7003)	(18.7861, 22.7080)

テストR二乗: 0.762701

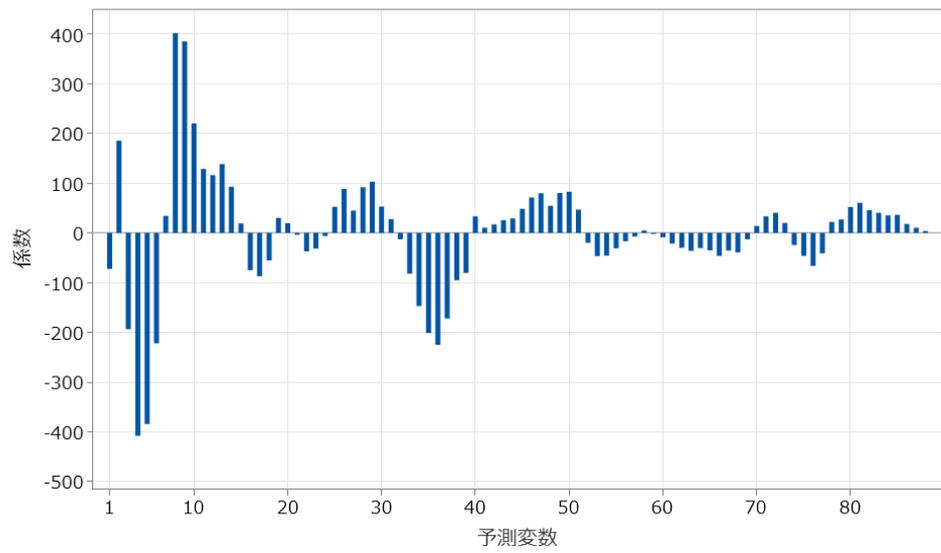
PLS係数プロット
(応答は水分)
10個の成分



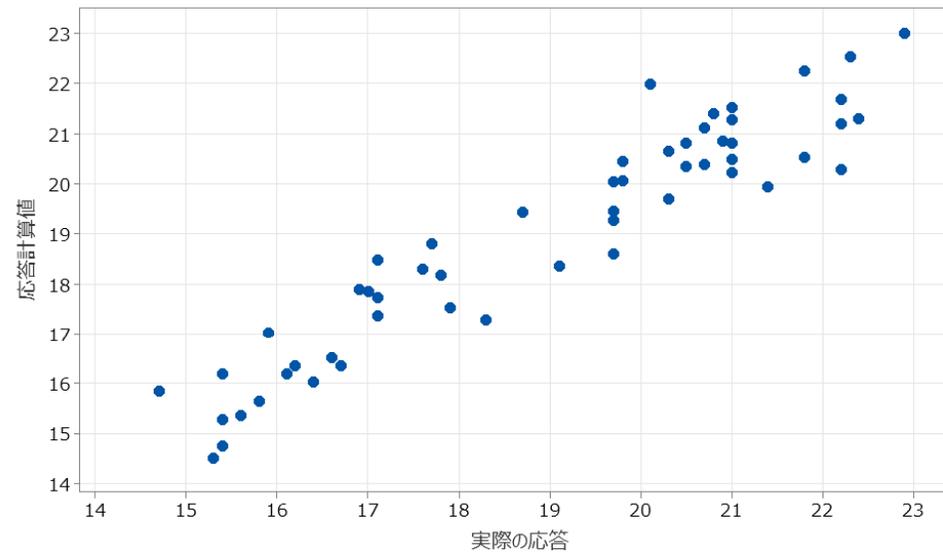
PLS応答プロット
(応答は水分)
10個の成分



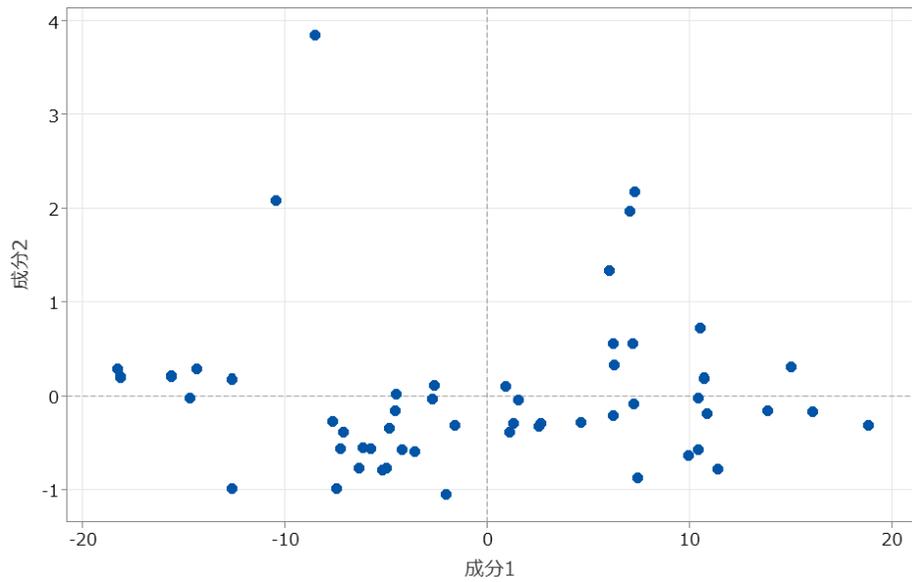
PLS係数プロット
(応答は脂肪)
10個の成分



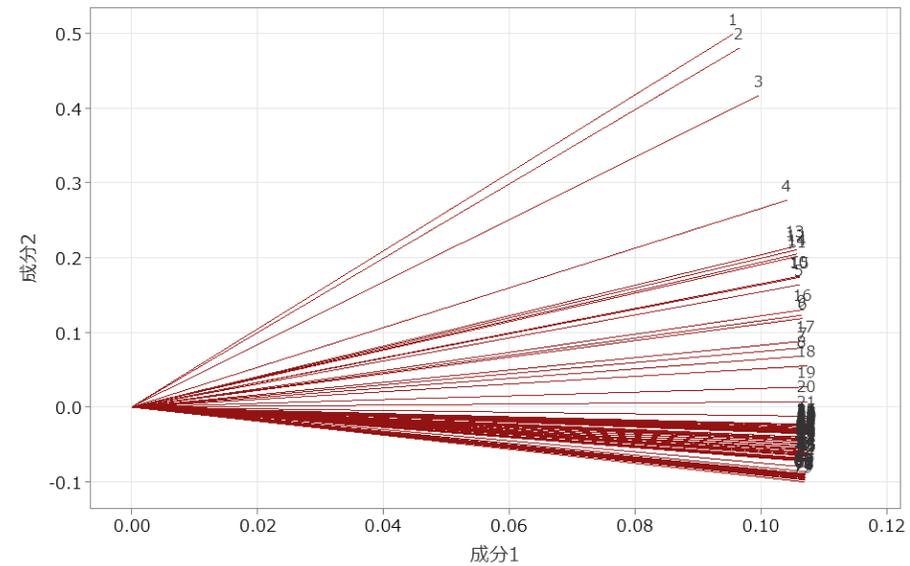
PLS応答プロット
(応答は脂肪)
10個の成分



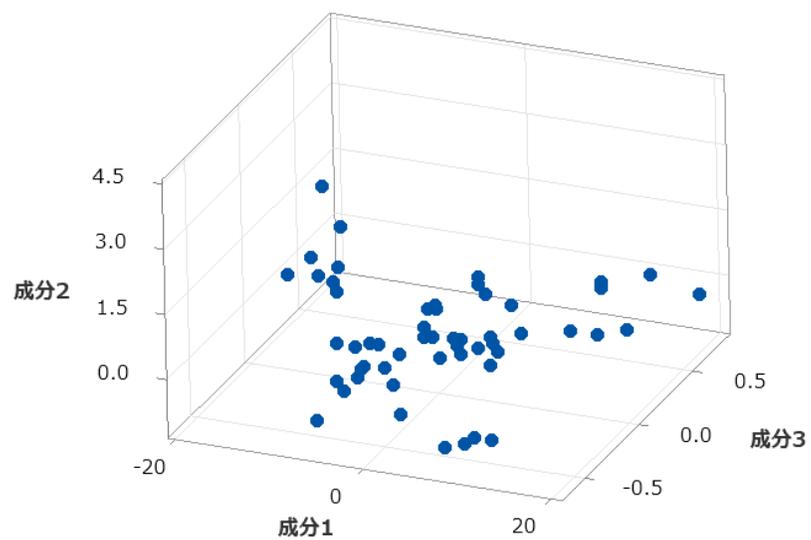
PLSスコアプロット



PLS因子負荷量プロット



PLS 3Dスコアプロット



PLSのX計算値のプロット
10個の成分

