できる Chemish



### **ChemInfoNavi** ケムインフォナビ

できるChemish 2004版R1

Copyright(C) 2004 ChemInfoNavi/CACフォーラム

http://www.cheminfonavi.co.jp/cac/

東京大学 船津公人研究室

製作 (有)ケムインフォナビ 441-8113 愛知県豊橋市西幸町字浜池333-9 豊橋サイエンスコア4F-401H Tel. 050-3315-2837

### できる Chemish 目次

第1章	はじめに	1
	Chemish とは	2
	サンプルデータ	3
第2章	基本的な操作方法・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・	5
	ファイルの新規作成	6
	ファイルの読み込み	7
	ファイルの保存	8
	シートの削除	9
	シートのコピー	1(
	シートの列選択	1
	シートの転置	12
	2 次元散布図	1:
	3 次元散布図	14
	折れ線グラフ	1!
	棒グラフ	16
	グラフの重ね合わせ	1
	グラフを読む	18
	グラフのコピー	19
	グラフの印刷	20
	グラフの保存	2
	グラフの読み込み	2
	グラフの設定	24
	画面の変更	2
	セルのカラー表示	2
第3章	基本的な統計手法・・・・・	2
	基礎統計量	28
	相関行列を計算する	29
	相関行列に色を付ける	30
	相関行列からの散布図の表示	3
	主成分分析の計算	32
	主成分分析の寄与率をグラフ化する	33
	主成分分析の T−T プロット	34
	重回帰分析	3
	Y-Ycalc プロット	3(
	回帰モデルを用いた予測	3.
	一括予測	38
	ステップワイズ変数選択	39

第4章	PLS 回帰、QPLS 回帰 · · · · · · · · · · · · · · · · · ·	41
	PLS 回帰分析	42
	R <sup>2</sup> のグラフを表示する	43
	R <sup>2</sup> 、Q <sup>2</sup> のグラフを表示する	44
	Y-Ycalc プロットを表示する	45
	PLS モデルを用いた予測	46
	GAPLS 法	47
第5章	クラスタリング ・・・・・・・・・・・・・・・・・・・・・・・・・・・・・	49
	階層型クラスタリングとは	50
	クラスタリング	51
	クラスタリング結果の表示	52
第6章	Back Propagation ニューラルネットワーク・・・・・・・・	53
	BP ニューラルネットワークとは	54
	BP ニューラルネットワークの学習	56
	学習回数-Q <sup>2</sup> プロットを表示する	58
	中間層の最適ユニット数を求める	59
	予測を行う	60
第7章	品質管理	61
	品質管理とは	62
	ヒストグラムの表示	63
	パレート図の表示	64
	管理図の表示	65
第8章	Kohonen ネットワーク、Counter-Propagation ·····	67
	Kohonen ネットワークとは	68
	Kohonen ネットワークの学習	69
	重みベクトルを表示する	71
	学習データを発火させる	72
	U-Matrix を見る	73
	CP とは	74
	CP の学習	75
	重みベクトルを表示する	77
	応答曲面を表示する	78
第9章	逆解析⋯⋯⋯⋯⋯⋯	79
	逆解析とは	80
	候補を提案する	81
	候補のスクリーニングを行う	84



# 第1章 はじめに

この章では、ケモメトリックスソフトウェアChemishの紹介を行いま す。また、この文書において用いられるサンプルデータについて説明 します。

この章の内容

● Chemishとは● サンプルデータ

## Chemishとは

#### Chemishとは

Chemishは、東京大学 船津公人研究室 (前豊橋技術科学大学) で開発されている、ケモメトリックスソフトウエアです。1998年にVer.1.0を発表して以来、毎年バージョンアップを重ね、現在も開発が続けられています。詳細については、

#### Chemishホームページ

http://www.cheminfonavi.co.jp/chemish/

#### をご覧下さい。

現在Chemishは使用できる機能の違いによっ てプロフェッショナルバージョンとスタンダー ドバージョンとがあり、後者をCACフォーラ ム会員が使用できます。

ChemishはMicrosoft Windows上で動作する ソフトウェアで、動作環境はWindows 95 / 98 / Me / NT / 2000 / XPです。



CACフォーラムとは

Computer Aided Chemistry Forum (CAC フォーラム)は、CHEMICS研究会の活動を 基に、実践分野の拡大を図るため、平成8年 4月に新たに再スタートした研究会です。構 造解析、合成設計およびケモメトリックスを 中心としたコンピューターケミストリーに関 する技術の、普及、取得及び開発研究を目指 しています。CACフォーラムは産学官の会員 から構成され、構造解析部会、合成設計部会 及びケモメトリックス部会の部会活動を中心 に運営されます。更に、CHEMICS研究会精 神を尊重し、各部会を有機的に連携させた全 体活動、会員のレベルを考慮した技術指導及 び特定テーマに限定したより実践的な分科会 活動も行っています。詳細については、 CACフォーラムホームページ

<u>http://www.cheminfonavi.co.jp/cac/</u> をご覧下さい。

#### 主な機能

Chemishは化学データの編集、解析、可視化の ための豊富な機能を備えています。以下に Chemishの主な機能を示します。

✓ 統計計算

- ✓ 平均、分散、標準偏差など
- ✓ 相関行列、主成分分析 (PCA)
- ✓ データモデリング
  - ✓ 最小2乗法による線形重回帰分析
  - ✓ PLS、QPLS
  - ✓ ニューラルネットワーク
- ✓ 変数選択
  - ✓ StepWise (変数増減 / 減増) 法
  - ✓ GAPLS、GAQPLS
- ✓ クラスタリング
  - ✓ 階層型クラスタリング
  - ✓ Kohonen ニューラルネットワーク
- ✓ 品質管理(QC)
  - ✓ ヒストグラム、パレート図、各種管理図
- 🗸 グラフ描画
  - ✓ 棒グラフ
  - ✓ 折れ線グラフ
  - ✓ 2次元散布図
  - ✓ 3次元散布図
  - ✓ 曲面グラフ

ケモメトリックスとは

ケモメトリックス (Chemometrics) とは chemistry と metrics を組み合わせた造語で、 日本語では計量化学と訳されます。統計学や 数学などの各種手法を化学データの解析に応 用し、有益な知見を得ることを目的とした学 問分野です。手法としては主成分分析、重回 帰分析、PLS、ニューラルネットワークなど が用いられ、構造一活性(物性)相関、材料 設計、品質管理などの幅広い分野に用いられ ています。

## サンプルデータ

### サンプルデータについて

本ドキュメントでは、Chemishの紹介のため のサンプルデータとして、主に2つのデータを 使用します。1つは "テストのデータ" であり、 PLSやBPなどのモデリング機能の紹介におい て使われます。もう1つは "黒曜石のデータ" で あり、モデリング以外の解析機能の例を示す際 に使用します。このページでは、これらのデー タについて説明をします。

#### 黒曜石のデータ

黒曜石のデータ (stone.csv) は、サンフラン シスコ湾付近の4ヶ所の採掘地で採取された黒 曜石について、蛍光X線を用いて微量元素の含 有量 (ppm) を測定したものです。測定された のは以下の10種類の元素です。

Fe, Ti, Ba, Ca, K, Mn, Rb, Sr, Y, Zr

各採掘地には1から4までの番号が対応付け られており、各行の先頭にその番号が付いてい ます。サンプル数は全部で63であり、その内 訳は以下のようになっています。

採掘地 1	10 個
採掘地 2	9個
採掘地 3	23 個
採掘地 4	21 個

各元素の含有量が行列に格納されており、各 行、各列にはラベルが付けられています。各列 のラベルは測定された元素の種類、各行のラベ ルはそのサンプルがどの採掘地で採取されたか を示す1から4までの数字です。また、各列の 平均が0、分散が1になるように規格化されて います。

このデータの出典を以下に示します。

佐々木愼一 他、化学者のためのパターン認識 序説、1984 テストのデータ

テストのデータ (test.csv) は、ある試験にお ける、各個人の点数と勉強時間などとの関係を 集計した仮想的なデータです。PLSや線形重回 帰分析などのモデリング手法の説明において使 用されます。

サンプルはA君からZ君までの26人です。目 的変数はその試験での点数、説明変数は試験前 の一週間における勉強時間、テレビを見た時間、 睡眠時間と前回の試験の点数、授業態度を5段 階で評価した数値の5つとなっています。



Chemishでは解析データを保存するための ファイル形式として、CSVとCHSを用いて います。CSV (Comma Separated Value)形 式とは、値をコンマで区切って保存する形式 です。Microsoft Excel などの表計算ソフトで 広く用いられています。そのため、それらの ソフトウェアを用いてChemishの入力データ を作成することや、Chemishの解析結果をそ れらのソフトウェアで読み込んで加工するこ となどが容易にできます。

CHS形式とはChemish独自のファイル形式 であり、複数の行列データを1つのファイル として保存することができます。これにより、 Chemishの解析結果をファイルに保存してお き、後でそのファイルを読み込んで参照する ことができます。





# 第2章 基本的な操作方法

この章では、ファイルの読み込みや保存、グラフの表示や読み方 など、Chemishを使用する上で必要となる基本的な操作法について 説明しています。

この章の内容

 ● ファイルの新規作成 ● ファイルの読み込み ● ファイルの保存 ● シートの削除 ● シートのコピー ● シートの列選択 ● シートの転置 ● 2 次元散布図 ● 3 次元散布図 ● 折れ線グラフ ● 棒グラフ ● グラフの重ね合わせ ● グラフを読む ●グラフのコピー ● グラフの印刷 ● グラフの保存 ● グラフの読み込み ● グラフの設定 ● 画面の変更 ● セルのカラー表示

# ファイルの新規作成







Chemish では、ファイルとシートの 2 つ の言葉を使用します。シートとは、Chemish 上でデータを表示するスプレッドシートのこ とを表します。Chemish では、計算の元に なるデータや、計算結果など複数のシートを 同時に扱います。ファイルとは単一、または 複数のシートをコンピュータのディスク上に 保存したもののことを言います。



シートは、データを記録したマスが2次元 に並んだもので、横方向を列、縦方向を行と 呼びます。列には説明変数や目的変数のデー タが、行にはサンプルデータが記録されます。 Chemish では、説明変数と目的変数を区別 するために、目的変数を黄色の背景で表示し ます。

ファイルの読み込み





ファイル"test.csv" が読み込まれ、
データシートが表示されます

🕽 Chemis	h - [test.cs	/]				
💁 ファイル	編集表示	テシート	計算 QC	グラフ ツー	ル ウインドウ	ハパ
D 🧀 🛛	MLR PCA	PLS RP	10 🕂 🛛	۳ 🖂 ۱	102	
w l					• 1 52 1 •	
AI	9424	=1.12	14417	***	-4-45 1997-0255 mile	12.84v
•	70000	7 VC	000.3C	100.000	仅未態度	
A D	70.000	20.000	50.000	F0.000	0.000	100.000
	20.000	30.000	49,000	00.000	3.000	100.000
0	45,000	20,000	42,000	96.000	5,000	00.000
E	54,000	12,000	42.000	77,000	1,000	90.000
F	40.000	33,000	30,000	75,000	5,000	95,000
6	46,000	20,000	38,000	99,000	3,000	90.000
н	31,000	8000	47.000	70.000	4 000	85,000
I	33,000	33,000	41.000	75.000	5,000	80.000
.l	47.000	6,000	47.000	50.000	2,000	70.000
ĸ	33,000	3000	44 000	55,000	3000	70.000
L	37.000	13.000	35.000	47.000	1.000	70.000
M	23.000	25.000	44.000	80.000	4.000	70.000
N	23.000	25.000	44.000	80.000	4.000	65.000
0	21.000	44.000	52.000	75.000	5.000	65.000
P	11.000	31.000	53.000	76.000	4.000	65.000
Q	20.000	10.000	46.000	55.000	3.000	65.000
R	16.000	33.000	47.000	46.000	3.000	50.000
S	20.000	41.000	55.000	48.000	4.000	50.000
Т	15.000	37.000	52.000	50.000	5.000	50.000
U	12.000	25.000	53.000	29.000	2.000	45.000
			10.000		1000	
Chemish for	Windows					



Chemish が扱うことのできるファイルは、 以下のようなものがあります。

- Chemish 専用ファイル
  CHS 形式 (\*.chs)
- ・データファイル
  - CSV 形式 (\*.csv)
- ・グラフファイル
  - Olectra Chart 2D 形式 (\*.oc2)
  - Olectra Chart 3D 形式 (\*.oc3)
- ・画像ファイル ・JPEG 形式 (\*.jpg)

これらの中で、シートに関係するものは、 CHS 形式と CSV 形式です。CHS 形式は、 計算の元になったデータの他に、計算結果な ど Chemish 上で扱われる全てのデータを保 存することができます。CSV 形式は多くの ソフトウェアで読み込むことできる特徴があ りますが、選択した一つのシートしか保存す ることができません。































種類 AdobePSGenericPostScriptPrinter 場所: FILE:	
אלאב	□ ファイルへ出力心
印刷範囲	印刷暗网数
<ul> <li>すべて(<u>A</u>)</li> </ul>	部数( <u>C</u> ): 1 🚞
○ ページ指定(G) 0 ページから(E)	
0 ページまで(T)	
○ 選択した部分(S)	14 24 34







用紙の幅、または高さに合わせて印刷が行 われます











Chapter			3	С	h		m			h
2	画面	面の	変更							
				_						
<u>1</u> メ=	ューの	選択			3	ステ	<mark>- ー タ ス</mark>	<mark>、バー</mark> 0	D非表示	Ę
Chemish - [Chemish2]     ステイル 編集 [表示 シー]     プール 編集 [表示 シー]     マール 新生 カラー     マール マール マール マール     マール マール     マール	甘茸 QC グラフ ツール     Ctrl+A     Ctrl+A	ウィンドウ ヘルブ 〇 <b>?</b> 現実数5 目的実数1 0.000	X _=0 x		تا خ	同様の操作 E非表示に	によって、 することも	ステーク できます	タスバー	
6 0000 00 7 0000 00 9 0000 00 9 0000 00 10 0000 00 10 0000 00 10 0000 00						Bit - [Chemish2]           イル 編集         表示         シー           120月次         カラー         カラー           1         0         ステータンパト           1         0         ステータンパト           1         0         ステータンパト           0         3         0.000         0.000           6         0.000         0.000         0.000           6         0.0000         0.000         0.000           9         0.0000         0.000         0.000           10         0.0000         0.000         0.000	計算 QC グラフ ツール Ctri+A 29月支払子 12년 0000 0 0000 0000 0 0000 0000	ウィンドウ ヘルブ 相支数5 目分支数1 0000	L X L A X	
<ol> <li>(1) メニュー または、 選択する</li> </ol>	・から、[表示 [表示] → [フ )	⊼] → [ツー ステータス	·ルバー] バー] を							
2 y-	ールバー	<mark>の非表</mark>	示		4	元に	戻す			
メニューの 消えたこの なくなりま アイル 編集 素示 2- マイル 編集 素示 2- マー 10 マステ央 4 000000	D [ステータ とにより、 ます + 計算 qc 377 ツール ctri+A 期刊まれ 調 パパー 0000 0000 00000	マスバー] ( ツールバー) ウルドウ ヘルブ 明定時5 日がまれ1 0000 0000 0000 0000 0000 0000	のチェック -が表示さ	がれ		hemish - [Chemish2] フィル 編集 表示 シー ゆ ↓ Hat ガラー マ マールズ 1以明花 マ ステーや 1 0000 0 3 0,000 0 3 0,000 0 5 0,000 0 5 0,000 0 7 0,000 0 8 0,000 0 9 0,000 0 10 0,000 0	ト 計算 QC グラフ ツール Ctrl+A レ □ □ ソト 取用欠数A 0000 0000 0000 0000 0000 0000 0000 0			
6 0.000 7 0.000 9 0.000 10 0.000 10 0.000 7ーム パーの 表示/非表示					<del>9-6</del>	<sup>ハーの表示/#表示</sup> 再度メニ で、ツーノ 示すること	ューにチェ ルバーやス こができま	ックを入 テータス す	れること バーを表	



2 カラー表示

セルの値が大きいものほど濃い赤で、	小さ
いまのにじ違い主ズ主ニキャキオ	
いものはと涙い月じ衣示されより	

😋 Cher	nish	- [stone.c	sv]							- 🗆 🗵
🔽 771	() L	編集表示	ネシート	計算 QC	グラフ ツー	ルーウィンドウ	フ ヘルプ		[	_ 8 ×
ID d	2 10		DIS DD	59 LL 18	e 🔽 🖌					
			100 01		<u>-</u>  ⊷ u	1 9 9				
XY										
	F	e	Ti	Ba	Ca	K	Mn	Rb	Sr	Y 🔺
	1	-0.336	1.161	0.750	0.858	1.232	-0.106	0.744	1.464	
	1	-0.110	1.403	0.690	1.006	0.884	0.033	1.635	1.348	
	1	-0.138	1.286	0.810	0.844	0.975	-0.314	0.744	1.521	
	1	-0.553	1.009	0.990	0.858	1.728	-0.592	1.219	1.233	
	1	-0.407	1.009	0.750	0.744	1.141	-0.592	-0.621	1.118	
	1	-0.398	1.277	0.630	0.855	0.902	-0.383	1.516	1.636	
	1	-0.584	0.893	0.990	0.726	1.471	-0.245	0.685	0.484	
	1	-0.491	1.215	0.810	0.873	1.636	0.102	1.932	2.443	
	1	-0.336	1.009	0.630	0.823	1.544	0.311	1.754	1.694	
	1	-0.432	1.027	0.510	0.995	0.664	-0.314	-0.443	1.118	
	2	-1.069	-0.691	-2.065	-0.197	1.085	-0.869	0.804	-0.955	
	2	-0.311	0.257	-2.124	0.367	0.609	-0.383	0.091	-0.955	
	2	0.004	0.141	-1.945	1.024	0.682	-0.175	0.566	-0.667	
	2	-0.011	0.275	-1.945	1.056	0.499	-0.245	0.448	-0.379	
	2	-0.336	0.060	-1.945	0.823	1.966	-0.453	2.229	-0.379	
	2	-0.336	0.177	-1.945	0.686	2.241	0.172	2.229	0.369	
	2	-1.607	-1.309	-2.005	-0.527	0.205	-1.424	0.151	-0.379	
	2	-0.070	-0.029	-1.945	0.930	0.700	-0.453	0.804	-0.667	
	2	-1.078	-0.700	-2.124	0.148	0.462	-0.939	0.448	-0.667	
	3	-0.794	-1.156	-0.328	-1.295	0.682	-0.869	0.626	-0.379	-
1	-									P
Chemish	for W	indows								



メニューのチェックが消えることで、通常 のセル表示が行われます

😋 Chemis	h - [stone.c	sv]							
🔕 ファイル	編集表示	テ シート	計算 QC	グラフ ツール	レ ウルドウ	ハバ			- 121 ×
D 🛋 🛛		PIS RD		• 🖌 m					
				•   • -   uo	<u>u                                    </u>				
XY									
	Fe	Ti	Ba	Ca	K	Mn	Rb	Sr	Y 🔺
1	-0.336	1.161	0.750	0.858	1.232	-0.106	0.744	1.464	
1	-0.110	1.403	0.690	1.006	0.884	0.033	1.635	1.348	_
1	-0.138	1.286	0.810	0.844	0.975	-0.314	0.744	1.521	
1	-0.553	1.009	0.990	0.858	1.728	-0.592	1.219	1.233	_
1	-0.407	1.009	0.750	0.744	1.141	-0.592	-0.621	1.118	
1	-0.398	1.277	0.630	0.855	0.902	-0.383	1.516	1.636	
1	-0.584	0.893	0.990	0.726	1.471	-0.245	0.685	0.484	
1	-0.491	1.215	0.810	0.873	1.636	0.102	1.932	2.443	
1	-0.336	1.009	0.630	0.823	1.544	0.311	1.754	1.694	
1	-0.432	1.027	0.510	0.995	0.664	-0.314	-0.443	1.118	
2	-1.069	-0.691	-2.065	-0.197	1.085	-0.869	0.804	-0.955	
2	-0.311	0.257	-2.124	0.367	0.609	-0.383	0.091	-0.955	
2	0.004	0.141	-1.945	1.024	0.682	-0.175	0.566	-0.667	
2	-0.011	0.275	-1.945	1.056	0.499	-0.245	0.448	-0.379	
2	-0.336	0.060	-1.945	0.823	1.966	-0.453	2.229	-0.379	
2	-0.336	0.177	-1.945	0.686	2.241	0.172	2.229	0.369	
2	-1.607	-1.309	-2.005	-0.527	0.205	-1.424	0.151	-0.379	
2	-0.070	-0.029	-1.945	0.930	0.700	-0.453	0.804	-0.667	
2	-1.078	-0.700	-2.124	0.148	0.462	-0.939	0.448	-0.667	
3	-0.794	-1.156	-0.328	-1.295	0.682	-0.869	0.626	-0.379	•
- ·									
Chemish for	Windows								/



### 第3章 基本的な統計手法

この章では、Chemish を用いて基本的な統計解析を行います。平 均値や分散などの基礎統計量の計算や、相関行列、主成分分析に よるデータ解析、重回帰分析による線形モデリングなどについて説 明します。

この章の内容

基礎統計量
 相関行列を計算する
 相関行列に色を付ける
 相関行列からの散布図の表示
 主成分分析の計算
 主成分分析の寄与率をグラフ化する
 主成分分析のT-Tプロット
 重回帰分析
 Y-Ycalcプロット
 回帰モデルを用いた予測
 ー括予測
 ステップワイズ変数選択

- 🗆 ×

\_ | # | ×

63( -25 21 0( 1( 1)

63.000 -2.164 2.229 -0.000 -0.000 1.000 1.000

63.000 -1.243 2.443 0.000 0.000 1.000 1.000

NUM



計算結果として追加されるシートの名前は 計算元のシートの名前にS()を付けたものに なります。例えば、"XY"という名前のシー トについて基礎統計量を計算すれば、その計 算結果は"S(XY)"という名前のシートに格納 されます。

予測 一括予測 スコア計算

-0.453 -0.939 -0.869

F

NUM

-0.02( -0.70) -0.70( -1.15( 侯補の)

-0.070 -1.078 -0.794

1

」 基礎統計量を計算します

# 相関行列を計算する



# 相関行列に色を付ける



3 色が付けられる



相関行列の評価

相関行列を計算する主な目的は、変数間の 関係を見ることです。相関行列の中のそれぞ れの値を見て、変数間の相関について確認を 行います。相関係数が1に近い値の場合正の 相関が強く、-1に近い場合負の相関が強いこ とになります。

例えば回帰分析を行う場合には、相関係数 の絶対値が1に近いような説明変数の組があ ると、モデルの予測性に悪影響を与えるため、 どちらか一方を除くなどの処置が必要です。 また、目的変数との相関が1に近いような説 明変数が存在すれば、重回帰分析を行う必要 はなく、単回帰で十分ということになります。

Chemish には値の大小によって行列に色 を付けて表示するという機能があり、これを 用いると相関行列の中から絶対値の大きい値 を探すことが容易になります。使われる色は 赤と青で、値が大きいほど赤く、小さいほど 青く表示されます。0に近い数値は白く表示 されます。





3計算を開始する

寄与率を見ることで、どの主成分までを用 いれば良いかが分かります。また、スコアを 見ることでデータの分布の様子が、ローディ ングを見ることで各主成分の特徴が分かりま す。

② 最大成分数を設定する


第3章 基本的な統計手法

# 主成分分析の T-T プロット



# 重回帰分析



フ	ァイル	を読み込み	らます	
こ	こでは	"test.csv"	を用いていま	す

Chemis	h - [test.cs	1					- 🗆 ×
A ファイル	編集表	モーシート	計算の	ガラフ ツー	ルーウィンドウ	1 11-7	_181 ×
	unitario de ca		1.945		10 54515		
L) 🗁 🛛	🚽 MLR PCA	PLS BP	🗄 🕂 🖡	1 🖌 🕺	] 🔏		
XY				-			
	勉強	テレビ	睡眠	前回の点券	授業態度	占数	<b></b>
A	70.000	0.000	35.000	100.000	5.000	100.000	
в	20.000	30.000	50.000	50.000	3.000	100.000	
5	55.000	7.000	43.000	98.000	3.000	100.000	
D	45.000	20.000	42.000	90.000	5.000	98.000	
-	54.000	12.000	40.000	77.000	1.000	95.000	
ř	40.000	33.000	39.000	75.000	5.000	95.000	
a	46.000	20.000	38.000	99.000	3.000	90.000	
н	31.000	8.000	47.000	70.000	4.000	85.000	
1	33.000	33.000	41.000	75.000	5.000	80.000	
J	47.000	6.000	47.000	50.000	2.000	70.000	
K	33.000	3.000	44.000	55.000	3.000	70.000	
L	37.000	13.000	35.000	47.000	1.000	70.000	
м	23.000	25.000	44.000	80.000	4.000	70.000	
N	23.000	25.000	44.000	80.000	4.000	65.000	
Э	21.000	44.000	52.000	75.000	5.000	65.000	
P	11.000	31.000	53.000	76.000	4.000	65.000	
Q	20.000	10.000	46.000	55.000	3.000	65.000	-
R	16.000	33.000	47.000	46.000	3.000	50.000	
S	20.000	41.000	55.000	48.000	4.000	50.000	
Т	15.000	37.000	52.000	50.000	5.000	50.000	
U	12.000	25.000	53.000	29.000	2.000	45.000	-
			10.000	07.000	1000	10,000	

2

重回帰分析を選択する
------------

1	[計算	] →	[MLR] <sup>;</sup>	を選	沢			
Chemish	ı - îtest.csv	1						X
C 7741	編集表示		計算 OC ガラフ	ツール ウィ	ッドウ	ヘルプ	- 17	X
			お井谷山島		É			
	el mur poa	PLS BP	本の現代に丁重					
XY			180401223		_			
	勉強	テレビ	Kohonen		1 4	気数		
A	70.000	0.000	CP		0	100.000		
В	20.000	30.000	発火		0	100.000		
C	55.000	7.000			-00	100.000		
D	45.000	20.000	MLR	Ctrl+G	10	98.000		
E	54.000	12.000	PCA	Ctrl+H	10	95.000		
F	40.000	33.000	PLS	Ctrl+J	10	95.000		
G	45.000	20.000	QPLS	Ctrl+K	0	90.000		
T I	22,000	22,000	BP	Ctrl+L	10	85.000		
1	47,000	6.000	クラスタリング		10	70,000		
U V	22,000	2,000			10	70.000		
h	37,000	13,000	GAPLS	Ctrl+B	10	70.000		
M	23,000	25.000	GAQPLS	Ctrl+N	ň	70.000		
N	23,000	25.000	STEPWISE	Ctrl+M	10	65.000		
0	21.000	44.000	子祖		bo	65.000		
P	11.000	31.000	11/21		00	65.000		
Q	20.000	10.000	16.17別		00	65.000		
R	16.000	33.000	スコパ計算		0	50.000		
S	20.000	41.000	逆解释析		0	50.000		
Т	15.000	37.000	候補のクラスない	5	0	50.000		
U	12.000	25.000	1751 1092 2 275 23 2	~	_00	45.000		-
線形重回帰会	分析を行いま	ţ					NUM	1.

	≣∔'	笛纟	土 耳	ま	压	言刃 -	<del>,</del>	z	
	<u> </u>	异亚	ロオ	50	狛庄	口心	9 1	С О	
	<b>`</b> Ó <del>h</del> n -	++	۲ <u>ـ</u> ۲.	Ŀ	とい記	+0 1			
──」 追加されたンートを選択し、									
	計管	結里	を確	認す	ろ				
	пі <del> 7</del> -	ጥሀ 不	C HE	DID 7	ω.				
		7							
Chem	nish - [te acsv]								- 0 ×
	And the state of t							100 100 100 100 100 100 100 100 100 100	
💁 771	ル 編え 表示	シート 計	算 QC グ	ラフ ツール	ウィンドウ	ヘルプ		[	_   <del>0</del>   ×
00 771 D 🖻	「ル編』を表示 F MLR PCA P	シート計 PLSBP 🔡	箕 qc ク 計井 図	'∋フ ツール   <mark>⊭</mark>   Щ	ウィンドウ <b>?</b>	ヘルプ		[	_ 121 ×
0 771 0 6	ル編え表示 FMLR PCA P VAT	シート 計 P <b>LS BP </b> 🖪 Ycalc F	算 QC グ 副 <b>ポ 区</b> 22   Par	'∋⊃ ツール   <mark>⊭</mark>   Щ   	ウィンドウ <b>?</b>	ヘルプ			_18 ×
0 771 0 2	ル編に、表示 ・ MLR PCA P ・ VAT 勉強 テ	シート 計 P <b>US BP   目</b> Ycalc   F Fレビ   間	算 QC グ    <b>ポ 区</b> 12   Par   眠    育	'∋フ ツール   <mark>▶</mark>   Щ         1回の点巻 携	ウィンドウ <b>?</b> 業態度	ヘルプ			- 181 ×
0 771 0 2	ル 編え 表示 ・ MLR PCA P ・ VAT 勉強 テ 70.000	シート 計 <b>213 BP 開</b> Yeale F Fレビ 順 0.000	算 QC グ 計 <u>ポ 区</u> 2 Par 画眼 育 35.000	「ラフ ツール ▶ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □	ウィンドウ ? 業態度 5.000	ヘルプ 点数 100.000			_ @ ×
Q 771 D 28 XY A B	「ル 編』を表示 ■ MER PCA P VAT 勉強 7 70.000 20.000	シート 計 <b>ILS BP 開</b> Yeale F Fレビ 増 0.000 30,000	算 QC グ 二十 区 2 Par 回眠 前 35.000 50.000	ラフ ツール ▶ □ 100.000 50.000	ウィンドウ <b>?</b> 業態度 5.000 3.000	ヘルプ 点数 100.000 100.000			_ @ ×
₿ 77-1 XY A B C	ル編え表示 ・ MR PGA P ・ VAT 勉強 700000 20000 55.000	シート 計 <b>2L3 BP 日</b> Yealc F Fレビ 増 0.000 30.000 7.000 7.000	算 QC グ 井 区 2 Par 回民 前 35.000 50.000 43.000	ラフ ツール ▶ ① 100.000 50.000 98.000 98.000	ウィンドウ <b>?</b> 業態度 5.000 3.000 3.000	ヘルプ 点数 100.000 100.000 100.000			_ @ ×
A B C D	ル 編集 表示 ・ MR PCA P ・ VAT - 20000 - 20000 - 55000 - 45000	シート 計 <b>PLS BP 開</b> Yeale F Fレビ 順 0.000 30.000 7.000 20.000 12000	算 QC グ 十 送 2 Par 回民 前 35.000 50.000 43.000 43.000 43.000	ラフ ツール ▶ □の点数 接 100.000 50.000 98.000 90.000 90.000 90.000	ウィンドウ <b>?</b> 業態度 5.000 3.000 3.000 1.000	ヘルプ 点数 100.000 100.000 100.000 98.000 95.000			_ @ ×
A B C D E	ル 編集 表示 MR PCA P VAT 勉強 20000 55.000 45.000 54.000 54.000	シート 計 ・ ・ ・ ・ ・ ・ ・ ・ ・ ・ ・ ・ ・	算 QC グ 計 総 2 Par 回版 前 35.000 43.000 43.000 42.000 40.000 20.000	ラフ ツール ▶ □ 100の点数 接 100.000 50.000 98.000 90.000 77.000 75.000	ウィンドウ <b>?</b> 業態度 5.000 3.000 3.000 5.000 1.000 5.000	ヘルプ 点数 100.000 100.000 98.000 95.000 95.000			_ 8 ×
A B C D E F	ル 編集 表示 MER POA P VAT <u>約時</u> 70.000 20.000 55.000 45.000 54.000 40.000	シート 計 <b>いのの</b> 「Vcalc F Fレビ 単 0.000 30.000 7.000 20.000 12.000 33.000 33.000 20.000	算 QC グ 2 Par 9度 前 35.000 43.000 43.000 43.000 43.000 39.000 39.000	ラフ ツール ▶ □ 100の点数 務 100.000 50.000 90.000 77.000 77.000 75.000 90.000	ウィンドウ <b>?</b> 業態度 5.000 3.000 3.000 5.000 1.000 5.000 3.000	ヘルプ 点数 100.000 100.000 98.000 95.000 95.000 95.000 95.000			<u> </u>
C 771 A B C D E F G H	ル 編え 表示 <b>FIR PCA P</b> <b>VAT</b> 物強 7 20000 55,000 45,000 40,000 40,000 40,000 31,000	シート 計 Yealc F Fレビ 第 0.000 30.000 7.000 20.000 12.000 33.000 20.000 8.000	算 QC 岁 文 2 Par 42 Par 42 Par 435.000 50.000 43.000 43.000 43.000 43.000 39.000 39.000 38.000 38.000	ラフ ツール ↓ ① 1000点形物 10000 50000 90.000 90.000 77.000 75.000 99.000 90.000	ウィンドウ <b>?</b> 業態度 5.000 3.000 5.000 1.000 5.000 3.000 4.000	ヘルプ 100.000 100.000 98.000 95.000 95.000 95.000 95.000			<u> </u>
A B C D E F G H H	アル 編集 表示 ・ ・ ・ ・ ・ ・ ・ ・ ・ ・	シート 計 Ycalc F Fレビ 第 0.000 30.000 20.000 12.000 33.000 20.000 33.000 20.000 33.000 33.000 33.000	算 QC 岁 2 Par 9 十 2 2 Par 9 1 35.000 50.000 43.000 42.000 40.000 39.000 38.000 47.000 41.000	ラフ ツール ↓ □の点表 推 100.000 50.000 98.000 90.000 77.000 75.000 99.000 75.000 99.000 75.000	ウインドウ <b>?</b>	ヘルプ 点数 100.000 100.000 98.000 95.000 95.000 90.000 85.000 80.000			
A B C D E F G H H I I	7ル 編集 表示 <b>148 PCA P</b> VAT 物強 7 70000 55000 55000 45000 45000 46000 46000 41000 31000 32000	シート 計 Ycalc F Fレビ 順 0.000 30.000 7.000 20.000 12.000 12.000 33.000 20.000 8.000 33.000 20.000	算 QC グ 日本 ビ 22 Par 時間 前 35.000 43.000 40.0000 40.00000 40.00000 40.0000 40.000000000 40.00000	ラフ ツール ▶ ① 1000点数 扱 100,000 50,000 98,000 98,000 98,000 99,000 75,000 75,000 75,000 50,000	ウインドウ <b> </b>	ヘルプ 点数 100.000 100.000 98.000 95.000 95.000 90.000 85.000 80.000 80.000 00.000			_ @ ×
A B C D E F G H J J K	アル 編集 表示 ・ 11R PGA P ・ 11R PGA P ・ 11R PGA P ・ 120000 20000 55000 45000 55000 45000 55000 40000 40000 40000 40000 31000 33000 33000	シート 計 <b>15 BP </b> 「Ycalc F FLピ ■ 0.000 30.000 12.000 12.000 33.000 80.000 33.000 80.000 33.000 80.000 33.000	算 QC グ	ラフ ツール ↓ □ 100.000 50.000 98.000 98.000 99.000 77.000 77.000 75.000 55.000	ウインドウ <b> </b>	ヘルプ 点数 100,000 100,000 98,000 95,000 90,000 80,000 70,000 70,000			_ @ ×
A B C D E F G H J J K L	アル 編集 表示 ・ ・ ・ ・ ・ ・ ・ ・ ・ ・	シート 計 Yealc F FLビ F FLビ F 20000 12000 12000 33000 20000 8000 33000 33000 33000 33000 12000 33000 12000	算         QC         グ           →         ↓         ↓           ●         →         ↓           ●         ↓         ↓	ラフ ツール ↓ ① 100.000 50.000 90.000 90.000 77.000 75.000 75.000 75.000 55.000 55.000 47.000	ウィンドウ <b>?</b> 実態度 5000 3000 5000 5000 4000 5000 2000 3000 1000	ヘルプ <u>点数</u> 100.000 100.000 96.000 95.000 95.000 95.000 85.000 80.000 70.000 70.000 70.000 70.000			
A A B C D E F G H I J K L L M	アル 編集 表示 ・ ・ ・ ・ ・ ・ ・ ・ ・ ・	シート 計 PLS BP E YCalc F Fレビ 単 0.000 30:000 20:000 20:000 20:000 8:000 3:0000 3:000 3:0000 3:0000 3:0000 3:0000 3:0000	算 QC グ 算 AF 22 22 Par 8000 43.000 43.000 43.000 43.000 43.000 43.000 44.000 38.000 44.000 44.000 44.000	ラフ ツール 第三の点表 排 100000 50000 90000 77,000 77,000 70,000 70,000 50,000 50,000 50,000 50,000 50,000 50,000 50,000 50,000 50,000 50,000 50,000 99,000 70,000 50,000 70,000 70,000 50,000 70,000 70,000 50,000 70,000 70,000 50,000 50,000 70,000 50,000 50,000 70,000 50,00	ウインドウ <b>?</b>	ヘルプ 点数 100.000 100.000 96.000 96.000 96.000 90.000 85.000 85.000 70.000 70.000 70.000 70.000			
A B C D E F G H J K L M N	プル 編集 表示     マル 編集 表示     マルス PGA P     マム P     マ	シート 計 *15 BP 「Ycalc F Fレビ 単 7,000 20,000 20,000 12,000 20,000 12,000 8,000 8,000 8,000 8,000 13,000 13,000 13,000 13,000 25,000	算         QC         グ           ++         上         L           22         Para         1           35.000         50.000         43.000           42.000         43.000         38.000           38.000         38.000         47.000           44.000         35.000         44.000           35.000         44.000         35.000	ラフ ツール レーシー 第回の点景 推 100,000 90,000 90,000 90,000 77,000 90,000 75,000 70,000 75,000 70,000 75,000 70,000 75,000 70,000 75,000 75,000 70,000 75,000 70,000 75,000 70,000 75,000 70,000 75,000 70,000 75,000 70,000 75,000 80,000 70,000 70,000 70,000 70,000 70,000 70,000 70,000 70,000 70,000 80,000 70,000 70,000 80,000 70,000 80,000 70,000 80,000 80,000 70,000 80,000 80,000 70,000 80,000 80,000 80,000 70,000 80,0	ウィンドウ <b>?</b>	ヘルプ 点数 100,000 100,000 96,000 95,000 95,000 95,000 80,000 80,000 70,000 70,000 70,000 65,000			
A B C D E F G H I J K L M N O	アル 編集 表示 アル 編集 表示 マロン のの マロン のの マロン のの ちものの ちまのの ちものの ちものの ちものの ちものの ちものの ちものの ちものの ちものの ちものの ちものの ちものの ちまのの ちものの ちものの ちものの ちものの ちものの ちものの ちものの ちものの ちものの ちものの ちまのの ちまのの ちまのの ちまのの ちまのの ちょの ちょ	シート 計	算 QC ク 算 → 上 2 Par 6度 第 35.000 50.000 50.000 43.000 43.000 43.000 43.000 44.000 47.000 47.000 44.000 44.000 44.000 52.000	ラフ ツール レー 第回の点番 携 100,000 50,000 98,000 99,000 99,000 77,000 99,000 77,000 99,000 77,000 99,000 50,000 50,000 50,000 50,000 80,000 80,000 80,000 77,500	ウィンドウ <b>?</b>	ヘルプ 100.000 100.000 100.000 98.000 98.000 96.000 90.000 80.000 70.000 70.000 70.000 70.000 70.000 65.000			
A B C D E F G H I J K L M N O P	パル 編集 表示     ペート     ペー     ペート     ペー	シート 計 *15 BP *16	算 QC グ 算 ++ 22 22 Par 第 35,000 43,000 43,000 40,000	ラフ ツール ↓ ↓ ↓ ↓ ↓ ↓ ↓ ↓ ↓ ↓ ↓ ↓ ↓ ↓ ↓ ↓ ↓ ↓ ↓	ウィンドウ <b>※</b>	ヘルプ 点数 100.000 100.000 95.000 95.000 95.000 85.000 88.000 70.000 70.000 70.000 65.000 65.000			
A B C D E F G H I J K L M N O P Q	アル 編集 表示 ・ ・ ・ ・ ・ ・ ・ ・ ・ ・ ・ ・ ・ ・ ・ ・ ・ ・ ・	シート 計	算 QC グ 第 十 22 2 Par 600 43,000 43,000 43,000 43,000 43,000 43,000 43,000 43,000 44,000 41,000 41,000 41,000 44,000 44,000 44,000 53,000 54,000	ラフ ツール シー (1) 「ロク」点様 推 「ロク」点様 「加 「ロク」点様 「加 」 「ロク」点様 「加 」 「ロク」点様 「加 」 「ロク」点様 「加 」 「ロク」点様 「加 」 「ロク」点様 「加 」 「ロク」点様 「加 」 「ロク」点様 「加 」 「ロク」点様 「加 」 「ロク」点様 「加 」 「ロク」」点様 「加 」 「ロク」点様 「加 」 「ロク」」点様 「加 」 「ロク」」点様 「加 」 「ロク」」」 「ロク」」」 「ロク」」 「ロク」」 「ロク」」 「ロク」」 「ロク」」 「ロク」」 「ロク」」 「ロク」」 「ロク」」 「ロー 」 「ロー 」 「ロー 」 「ロー 」 「ロー 」 「ロー 」 「ロー 」 「ロー 」 「 」 」 「 」 」 「 」 」 」 「 」 」 」 「 」 」 」 「 」 」 」 」 」 「 」 」 」 」 」 」 「 」 」 」 「 」 」 」 」 「 」 」 」 」 「 」 」 」 」 」 」 「 」 」 」 」 」 「 」 」 」 「 」 」 」 」 」 「 」 」 」 」 」 」 「 」 」 」 」 」 」 」 」 」 」 」 「 」 」 」 」 」 」 」 」 」 」 」 」 」	ウィンドウ <b>?</b>	<u> </u>			
A B C D E F G H I J K L M N O P Q R	ル 編 表示 ・ ・ ・ ・ ・ ・ ・ ・ ・ ・ ・ ・ ・ ・ ・ ・ ・ ・ ・	シート 計 Ycalc F FLC 単 0000 30000 7000 12000 12000 30000 20000 12000 30000 12000 30000 12000 30000 12000 30000 13000 1	算         QC         グ           1         1         1         1           12         Par         1         1         1           12         Par         1         1         1         1           13         50.000         1         35.000         1	ラフ ツール レーロー 第回の点載 排 10000 50,000 90,000 90,000 77,000 77,000 77,000 77,000 75,000 50,000 50,000 50,000 55,000 47,000 75,000	ウィンドウ <b>※</b>	ヘルプ 点数 100,000 100,000 96,000 96,000 96,000 80,000 70,000 70,000 70,000 70,000 65,000 65,000 65,000 65,000			
ABCDEFGHIJKLMNOPQRS	<sup>1</sup> ル 線体 表示 <sup>1</sup> 内線 表示 <sup>1</sup> 内線 表示 <sup>1</sup> 内線 次和 <sup>1</sup> マンコーン <sup>1</sup> マンコー	シート 計 ************************************	算         QC         グ           日         日         日         12           1         1         12         12           1         1         12         12           1         1         13         13           1         30.000         14         1000           1         39.000         14         1000           1         39.000         14         1000           1         10.000         15         1000           1         10.000         14         1000           1         10.000         14         1000           1         10.000         15         1000           1         10.000         14         1000           1         10.000         14         10.000           1         10.000         15         10.000           1         10.000         14         10.000           1         10.000         14         10.000           1         10.000         14         10.000           1         10.000         14         10.000           1         10.000         14         10.000 <t< td=""><td>ラフ ツール ・ 「日の大美様 様 100,000 98,000 90,000</td><td>ウィンドウ <b>家</b></td><td>ヘルプ 点数 100.000 100.000 98.000 95.000 95.000 95.000 70.000 70.000 70.000 70.000 65.000 65.000 65.000 50.000</td><td></td><td></td><td></td></t<>	ラフ ツール ・ 「日の大美様 様 100,000 98,000 90,000	ウィンドウ <b>家</b>	ヘルプ 点数 100.000 100.000 98.000 95.000 95.000 95.000 70.000 70.000 70.000 70.000 65.000 65.000 65.000 50.000			
ABCDEFGHIJKLMNOPQRST	10 日本 表示 10 日本 表示 10 日本 次日 10 日本 次日 10 日本 10 日本 1	シート 計      SP      SP      Ycalc     F      Ycalc	算         QC         グ           1         1         1         1           12         Par         1         1         1           13         0.00         1         1         1         1           14         0.00         1         <	ラフ ツール ↓ □ 100,000 50,000 90,000 77,000 90,000 77,000 90,000 77,000 50,000 50,000 50,000 50,000 76,000 50,000 76,00	ウィンドウ <b>学</b>	<ul> <li>ヘルプ</li> <li>点数</li> <li>100,000</li> <li>100,000</li> <li>96,000</li> <li>96,000</li> <li>96,000</li> <li>80,000</li> <li>80,000</li> <li>80,000</li> <li>70,000</li> <li>70,000</li> <li>70,000</li> <li>70,000</li> <li>66,000</li> <li>65,000</li> <li>65,000</li> <li>50,000</li> <li>50,000</li> <li>50,000</li> <li>50,000</li> </ul>			_ 8 ×



重回帰分析 (Multi-Linear Regression)とは、 複数の説明変数からなるXと目的変数yとの 間で回帰式を作成して解析を行う手法です。 線形重回帰分析ではモデル式は、

y = Xb + e

となります。ここでbは回帰係数、eは残差を 表しています。Chemishでは最小2乗法を用 いて、残差の2乗和が最小になるようなモデ ル式を構築することができます。







Chapter <sup>c</sup> <sup>e</sup> <sup>o</sup> <sup>c</sup>	n e m i s n
3 一括予測	
1 重回帰分析を行う	3 予測用ファイルを開く
重回帰分析を行います	ここでは "test_pred.csv" を用います
Chernich - [test.csv]     スケイル 編集 表示 シート 計算 QC グラフ ツール ウインドウ ヘルフ (メータ)     スケイル 編集 表示 シート 計算 QC グラフ ツール ウインドウ ヘルフ (メータ)     スケイル 編集 表示 シート 計算 QC グラフ ツール ウインドウ ヘルフ (メータ)     スケイル 編集 表示 シート 計算 QC グラフ リール ウインドウ ヘルフ (メータ)     スケイル 編集 表示 シート 計算 QC グラフ リール ウインドウ ヘルフ (メータ)     スケイル 編集 表示 シート 計算 QC グラフ リール ウインドウ ヘルフ (メータ)     スケイル 編集 表示 シート 計算 QC グラフ リール ウインドウ ヘルフ (メータ)     スケイル 編集 表示 シート 計算 QC グラフ リール ウインドウ ヘルフ (メータ)     スケイル 編集 表示 シート 計算 QC グラフ リール ウインドウ ヘルフ (メータ)     スケイル (メーター)     スケイル     スケイル (メーター)     スケイル (メー	子源データファイルを開く       ?         ファイルの場所の
2 一括予測を選択	4 予測結果を確認する
O fitig doubleO fitig double	① 追加されたシートを選択し、 予測結果を確認する          ● Chemoid-Icest.cv/ 予測結果を確認する         ● アイル 編集 表示 シート 計算 cc グ57 ツール マンドウ ヘルプ         ● アイル 編集 表示 シート 計算 cc グ57 ツール マンドウ ヘルプ         ● アイル 編集 表示 シート 計算 cc グ57 ツール マンドウ ヘルプ         ● アイル 編集 表示 シート 計算 cc グ57 ツール マンドウ ヘルプ         ● アイル 編集 表示 シート 計算 cc グ57 ツール マンドウ ヘルプ         ● アイル 編集 表示 シート 計算 cc グ57 ツール マンドウ ヘルプ         ● アイル (abs BP B) デビビレ )         ● アイル (abs BP B) デビビレ )         ● アイル (abs BP B) デビビレ )         ● 17321 42488         ● 17321 42488         ● 17321 42488         ● 17321 42488         ● 17321 42488         ● 17321 42488         ● 17321 42488         ● 17321 42488         ● 17321 42488         ● 17321 4288         ● 17321 4288         ● 17321 4288         ● 17321 4288         ● 17321 4288         ● 17321 4288         ● 17321 4288         ● 17321 4288         ● 17321 4288         ● 17321 4288         ● 17321 4288         ● 17321 4288         ● 17321 4288         ● 17321 4288         ● 17321 4288         ● 17321 4288         ● 17188         ● 17388         ● 17188
② [一括予測] をクリック	



ステップワイズ変数選択



**2** 設定を行う







線形重回帰分析は、変数の数が多ければ多 いほど決定係数 (R<sup>2</sup>)の値が高くなるという 性質を持っています。しかし、それは偶然の 相関である可能性もあるため、必ずしも予測 的なモデルができているとは言えません。目 的変数に本質的に関係のない変数やノイズの 多く含まれている変数はモデルに入れるべき ではありません。

ステップワイズ変数選択法(逐次選択法) は、予測的なモデルを構築するために適切な 変数を自動的に決定する変数選択手法です。 増減法では、目的変数と最も相関の強い1つ の変数から始め、1つずつ変数を加えていき ます。逆に減増法では、すべての変数を用い たモデルから始め、不必要と思われる変数を 取り除いていき、最適な変数の組み合わせを 求めます。変数を追加あるいは削除するとき には通常F値を基準とします。Chemishで設 定可能なFin、Foutの値はこのときの閾値で す。



# 第4章 PLS回帰、QPLS回帰

この章では、PLS法、GAPLS法について紹介します。PLS法は、ケ モメトリックスの分野などで幅広く用いられている線形回帰分析手法 です。通常の最少2乗法による線形重回帰分析に比べ、予測的な回 帰モデルが得られることが知られています。また、GAPLS法は遺伝 的アルゴリズム (GA) を用いてPLS法に用いる変数を選択するため の手法です。

この章の内容

● PLS回帰分析
 ● R<sup>2</sup>のグラフを表示する
 ● R<sup>2</sup>、Q<sup>2</sup>のグラフを表示する
 ● Y-Ycalcプロットを表示する
 ● PLSモデルを用いた予測
 ● GAPLS法

	[計位]	1 <u> </u>	DI 01	た遅せ	P				
U	「可予」	ו – ו	FLOJ	で迭れ					
( <b>A</b> ) -1			California and a second				(Constantion)		1.5
C Lhen	nish - [test.csv]			- 24 - 44 - <del>44</del>	) Inde				
<b>GA</b> 771	ル編集表示	2-1	計算 QC クラ	9-10 94	210	NID.			212
] 🗅 🐸	🖬 MLR PCA F	PLS BP	基礎統計量 *9期約270						
XY			18(A) 129						
	勉強 ラ	FVĽ	Kohonen		-	皷	_		
A	勉強 5 70.000	Fレビ 0.00(	Kohonen CP		10	長数 100.000	_	_	ľ
A B	勉強 5 70.000 20.000	Fレビ 0.00( 30.00(	Kohonen CP 発火		10	<数 100.000 100.000			4
A B C	勉強 7 70.000 20.000 55.000	Fレビ 0.00( 30.00( 7.00(	Kohonen CP 発火	outro	10	5.5t 100.000 100.000 100.000			ľ
A B C D	<u>税残</u> 業 7 70.000 20.000 55.000 45.000	Fレビ 0.00( 30.00( 7.00( 20.00(	Kohonen CP 発火 MLR	Ctrl+G		5.新t 100.000 100.000 100.000 98.000			ľ
A B C D E	税3集 7 70.000 20.000 55.000 45.000 54.000 40.000	Fレビ 0.00( 30.00( 7.00( 20.00( 12.00( 22.00(	Kohonen CP 発火 MLR PCA	Ctrl+G Ctrl+H		5.長女 100.000 100.000 100.000 98.000 95.000 05.000			ľ
A B C D F	<u>税3</u> 建 70,000 20,000 55,000 45,000 54,000 40,000	0.00( 30.00( 7.00( 20.00( 12.00( 33.00( 20.00(	Kohonen CP 発火 MLR PCA PLS	Ctrl+G Ctrl+H Ctrl+J		5.50 100.000 100.000 98.000 95.000 95.000 90.000			1
A B C D F G H	<u>税33</u> 年 70.000 20.000 55.000 45.000 54.000 40.000 46.000 31.000	0.00( 30.00( 7.00( 20.00( 12.00( 33.00( 20.00( 8.00(	Kohonen CP 発火 MLR PCA PLS QPLS	Ctrl+G Ctrl+H Ctrl+J Ctrl+K		5.50 100.000 100.000 98.000 95.000 95.000 95.000 90.000 85.000			ŕ
A B C D E F G H	焼き	Fレビ 0.00( 30.00( 7.00( 20.00( 12.00( 33.00( 20.00( 8.00( 33.00( 33.00(	Kohonen CP 発火 MLR PCA PLS QPLS BP	Ctrl+G Ctrl+H Ctrl+J Ctrl+K Ctrl+K Ctrl+L		5.数 100.000 100.000 98.000 95.000 95.000 95.000 90.000 85.000 80.000			1
A B C D F G H I J	焼き	FUE 0.00( 30.00( 7.00( 20.00( 12.00( 33.00( 8.00( 33.00( 8.00( 33.00( 8.00(	Kohonen CP 発火 MLR PCA PLS QPLS BP クラスタリング	Ctrl+G Ctrl+H Ctrl+J Ctrl+K Ctrl+L		(数 100.000 100.000 98.000 95.000 95.000 95.000 90.000 85.000 80.000 70.000			1
A B C D E F G H J K	税会         ラ           20.000         20.000           55.000         45.000           54.000         54.000           40.000         46.000           31.000         33.000           47.000         33.000	FUE 0.00( 30.00( 7.00( 20.00( 12.00( 33.00( 8.00( 33.00( 3.00(	Kohonen CP 発火 MLR PCA PLS QPLS BP クラスタリング	Ctrl+G Ctrl+H Ctrl+H Ctrl+J Ctrl+K Ctrl+K		5.50 100.000 100.000 98.000 95.000 95.000 90.000 85.000 85.000 70.000 70.000			ŕ
A B C D E F G H I J K L	労会員 70,000     20,000     55,000     45,000     45,000     46,000     46,000     46,000     46,000     31,000     33,000     47,000     33,000     37,000     37,000	FLE 0.00( 30.00( 7.00( 20.00( 12.00( 33.00( 33.00( 33.00( 33.00( 30.	Kohonen CP 発火 MLR PCA PLS QPLS BP クラスタリング GAPLS	Ctrl+G Ctrl+H Ctrl+J Ctrl+K Ctrl+L Ctrl+B		5.52 100.000 100.000 98.000 95.000 95.000 90.000 85.000 80.000 70.000 70.000			ſ
A B C D E F G H J J K L	<u>約録</u> 7 20000 55.000 45.000 445.000 440.000 440.000 33.000 33.000 33.000 37.000 33.000 33.000 33.000 33.000 33.000 33.000	FUE 0.00( 30.00( 7.00( 20.00( 12.00( 33.00( 30.00( 30.00( 30.00( 30.00( 30.00( 25.00(	Kohonen CP 発火 MLR PCA PLS BP クラスがリング GAPLS GAQPLS	Ctrl+G Ctrl+H Ctrl+J Ctrl+K Ctrl+L Ctrl+B Ctrl+B		b           100.000           100.000           100.000           98.000           95.000           95.000           96.000           96.000           96.000           90.000           85.000           80.000           70.000           70.000           70.000           70.000			ľ
A B C D E F G H J K L M N	*的注意 ラ 70000 20000 55000 45000 46000 46000 46000 33000 33000 33000 37000 23000 23000     2300	FLE 0.00( 30.00( 7.001 20.00( 33.000 20.00( 8.00( 33.00( 33.00( 13.00( 25.00(	Kohonen CP 発火 MLR PCA PLS QPLS BP クラスタリング GAPLS GAQPLS STEPWISE	Ctrl+G Ctrl+H Ctrl+H Ctrl+K Ctrl+K Ctrl+L Ctrl+B Ctrl+N Ctrl+N		Ebb           100.000           100.000           100.000           98.000           95.000           95.000           95.000           95.000           95.000           95.000           95.000           95.000           95.000           95.000           95.000           95.000           70.000           70.000           70.000           65.000			ľ
A B C D E F G H I J K L M N O	★的会 ライロの (1) (1) (2) (2) (2) (2) (2) (2) (2) (2) (2) (2	FLE 0.00( 30.00( 7.00( 20.00( 12.00( 33.00( 8.00( 33.00( 6.00( 3.00( 13.00( 25.00( 44.00) 44.00( 10.00(	Kohonen CP 発火 MLR PCA QPLS BP クラスタリング GAPLS GAQPLS STEPWISE ぞぞ明	Ctrl+G Ctrl+H Ctrl+J Ctrl+K Ctrl+K Ctrl+R Ctrl+N Ctrl+M		kk           100.000           100.000           96.000           95.000           95.000           95.000           90.000           80.000           70.000           70.000           70.000           65.000           65.000			ľ
A B C D E F G H I J K L M N O P	* 第25章 70000 20000 55000 45000 45000 46000 46000 46000 46000 33000 33000 37000 23000 23000 23000 21000 11.000	LL"           0.000           30.000           7.000           20.000           12.000           33.000           20.000           31.000	Kohonen CP 発火 MLR PCA PLS QPLS BP クラスがリング GAPLS GAQPLS STEPWISE 子別 - ビーズの町	Ctrl+G Ctrl+H Ctrl+J Ctrl+K Ctrl+L Ctrl+B Ctrl+N Ctrl+M		Eby           100.000           100.000           98.000           95.000           95.000           90.000           80.000           80.000           70.000           70.000           70.000           65.000           65.000           65.000			ľ
A B C D E F G H I J K L M N O P Q	地話         7           70000         20000           55000         45000           45000         46000           31,000         33,000           47,000         33,000           23,000         23,000           21,000         11,000           20,000         21,000	FLE*           0.000           30.000           7.000           20.000           12.000           33.000           20.000           8.000           33.000           33.000           33.000           33.000           33.000           13.000           25.000           25.000           25.000           31.000           10.000	Kohonen CP 発火 MLR PCA PLS QPLS BP クラスタリング GAPLS GAPLS GAPLS STEPWISE 子川 一括子別別	Ctrl+G Ctrl+J Ctrl+J Ctrl+Z Ctrl+L Ctrl+B Ctrl+N Ctrl+M		Lab           100.000           100.000           100.000           98.000           95.000           96.000           96.000           96.000           96.000           96.000           96.000           96.000           96.000           96.000           70.000           70.000           70.000           65.000           65.000           65.000			
ABCDEFGHIJKJMNOPGR	地路         7           70.000         20.000           55.000         45.000           45.000         40.000           40.000         31.000           33.000         33.000           23.000         23.000           21.000         21.000           11.000         20.000	ELE"           0.000           30.001           7.001           20.000           33.001           20.001           33.001           8.001           33.001           6.001           30.001           25.001           25.001           25.001           31.001           25.001           33.001           33.001           33.001           33.001	Kohonen CP 発火 MLR PCA PLS QPLS BP クラスタリング GAPLS GAQPLS STEPWISE 子期 へ田子計算	Ctrl+G Ctrl+J Ctrl+J Ctrl+J Ctrl+J Ctrl+B Ctrl+N Ctrl+M		EBY           100.000           100.000           100.000           98.000           95.000           95.000           95.000           95.000           95.000           90.000           85.000           80.000           70.000           70.000           70.000           65.000           65.000           65.000           65.000           65.000			
A B C D H F G H I J X J M Z O P G R Ø	№№         7           70.000         20.000           55.000         45.000           45.000         46.000           31.000         33.000           33.000         33.000           23.000         23.000           21.000         11.000           16.000         16.000	FLE"           0.000           30.000           7.000           20.000           12.000           33.000           6.000           30.000           13.000           25.000           25.000           31.000           31.000           33.000           44.000           31.000           34.000           34.000           34.000           34.000           34.000	Kohonen CP 発火 MLR PCA PLS QPLS GAPLS GAQPLS STEPWISE 子測 一括予測 入口7計算 详範紙	Сагнна Сагння Сагння Сагня Сагня Сагня Сагня Сагня		Lab           100.000           100.000           100.000           100.000           98.000           95.000           95.000           90.000           80.000           70.000           70.000           70.000           70.000           70.000           65.000           65.000           65.000           50.000           50.000			

設定を行う

2





PLS法とは

PLS (Partial Least Squares regression)法 とはケモメトリックスの分野で標準的に用い られている線形回帰分析手法です。主成分分 析と同じように、変数からスコアを計算して モデリングを行います。説明変数Xからt、目 的変数yからuと呼ばれるスコアを抽出し、そ れらの相関が最大となるようにモデリングし ます。通常の最少2乗法による回帰に比べ、 頑健で予測的なモデルが得られることが知ら れています。また、サンプルの数よりも変数 の数の方が多い場合や、変数間に強い相関関 係が存在する場合においても、モデリングを 行うことが可能です。

モデルの評価は主にR<sup>2</sup>とQ<sup>2</sup>の値で行いま す。最大値はともに1で、値が大きいほど良 いモデルであることを示します。R<sup>2</sup>は通常の 決定係数、Q<sup>2</sup>は leave-one-out によるクロス バリデーションを行った場合の決定係数です。 成分数ごとにこれらの値を評価し、最適な成 分数を決める必要があります。







横軸が成分数、縦軸がR<sup>2</sup>の値です。 R<sup>2</sup>は成分数に対し、単調に増加します

R<sup>2</sup>、Q<sup>2</sup>のグラフを表示する



2

系列を追加する

グラフ設定			X
系列1 🚽			
Xee XY	▼ 勉強		~
Yeek R2	▼ 点数		•
Ziện XY	▼ 勉強		~
5	ラフの種類 折れ線	パラフ	•
□ データラベルの表示	True	eTypeFontを使用	
		系列の色	
系列の追加系列の	前除 🗌 25	スタリング情報のま	र्फ
	Ē	閉じる 🛛 🛄	עםכ
① [系列の追加]	]をクリッ	ク	

3 グラフの設定を行う

4 グ

グラフが表示される



横軸が成分数、縦軸がQ<sup>2</sup>の値です。 Q<sup>2</sup>は成分数に対し、単調には増加しません。 最適な成分数は、Q<sup>2</sup>が最大となる3であると 考えられます



😋 2次元グラフ

Ycalc [ 点数(1) ]

グラフ ヘルプ

120 110 100

> 90 80

> > 70

60

50

40

30

•



グラフを表示する

横軸がモデリングに用いた目的変数の値、縦軸 がPLSモデルによる計算値です。ほぼすべてのサ ンプルが対角線付近にあり、良好なモデルであ ることが確認できます

50

XY[点数]

A (100.000, 114.323)

60 70

30 40





2

<u>- 🗆 ×</u>

80 90 100

۲





GAPLSを選択

1

<b>C</b>	hemist	n - [tes	st.csv									<u>_     ×</u>
0	ファイル	編集	表示	シート	計算	QC	グラフ	ツール	ウル	ドウ	ヘルプ	_ 8 ×
	i 🖉 🖥	d mu	PCA	PLS BP	基	礎統言	+量					
1 XY					相称	期行列	a)					
P <sup>A</sup>		新己名	1		Koł	юпеп				F	占新	
A		70	0000	0.000	CP					10	100.000	-
в		20	0000	30.000	*	k				00	100.000	
C		55	5.000	7.000		· · ·				00	100.000	
D		45	5.000	20.000	MLF	۶		Ctrl	+G	00	98.000	
E		54	1.000	12.000	PCA	4		Ctrl	+H	0	95.000	
F		40	0.000	33.000	PLS			Ctrl	+J	00	95.000	
G		46	6.000	20.000	OP	LS		Ctrl	+К	00	90.000	
H		31	.000	8.000	BP			Chrl	+1	00	85.000	
I		33	3.000	33.000	5- 21-	7.60	- 18	Cur		00	80.000	
J		47	7.000	6.000		×33	<i>)</i> ,,			00	70.000	
K.		33	3.000	3.000	GA	PLS		Ctrl	+8	10	70.000	
L.		3	7.000	13.000	GA	OPLS		Ctrl	+N	10	70.000	
M		23	3.000	25.000	STE	PWT	F	Chrl	+M	10	70.000	
N		23	3.000	25.000						10	65.000	
He-		21	.000	44.000	- 予)						65.000	
HP-		00	.000	31.000		舌子湯				6	65,000	
		20	0000	22,000						6	50.000 E0.000	
le-		20	0000	41.000						6	50,000	
1 T		15	0000	97.000	逆	前折				6	50,000	
Hi-		13	2000	25,000	候	雨のク		j ji		10	45,000	
⊩≚–				20.000							40.000	•
GAT	変数運	択を行い	ます(	PLS)								NUM
						_		_		_		
	~ .			_								
(	1)	[ <u>=</u> +	窅	$1 \rightarrow$	[G	ΔI	21 3	SI :	を	望	択	
	Ч		᠇	J ^	Ľ	/ \l	_ L '	<b>_</b>	رے	25	. JV C	

設定を行う

2

① GAの各種パラメ·	ータを設定します
計算条件設定	×
▼ センタリング	-GAの設定
☑ スケーリング	100 👤 世代数
▼ クロスバレデーション	30 👤 染色体数
	30 💌 カットオフ
	30   生存率
5 _ 成分数	5 文美対数
1 <u>▼</u> 目的変数の数	
計算開始 キャンセル	
② 計算を開始します	-
<ul> <li>□ リスケーリング</li> <li>5 、 成分数</li> <li>1 、 目的変数の数</li> <li>計算関始 キャンセル</li> <li>② 計算を開始します</li> </ul>	30 30 30 ↓ 生存率 5 文差対数 1 マ 突然変異率



GAPLS法とは

GA (遺伝的アルゴリズム) とは生物の遺伝 の様子を模倣した最適化手法です。0と1で表 現された遺伝子に対し、突然変異や交差といっ た操作を行い、新たな遺伝子を作り出します。 そして各遺伝子について評価値を計算し、淘 汰、選択を行います。これによって、優れた 個体の周辺の空間が優先的に探索され、結果 として最適に近い解が効率よく発見できます。 アルゴリズムが簡単であり、また最適化能力 も比較的高いため、様々な分野において幅広 く用いられています。

GAPLS法とは、そのGAを用いた変数選択 手法です。遺伝子の各ビットに説明変数の各 変数を割り当て、最適なPLSモデルを与える 変数の組を探索します。遺伝子の評価関数と しては、leave-one-out 法によるクロスバリ デーションの結果であるQ<sup>2</sup>値を用います。こ れにより、予測的なモデルを構築することの できる変数の組み合わせが得られます。GA では初期条件などによって探索結果が変わる 可能性があるため、条件を変えて何度か計算 を行う必要があります。



# 第5章 クラスタリング

この章では、異なる性質を持ったサンプルをその類似度により分類 するクラスタリングと、Chemish でクラスタリングを行う方法について 説明します。

この章の内容

● 階層型クラスタリングとは
 ● クラスタリング
 ● クラスタリング結果の表示



階層型クラスタリング

階層型クラスタリングはクラスタリングの代 表的な手法で、以下のアルゴリズムに従ってサ ンプルの分類を行います。

- 1. それぞれ 1 つのサンプルのみを構成要素 とする、n 個のクラスタ (集団) を作る。
- 2. 各クラスタ間の類似度を比較し、最も類 似したクラスタを統合する。
- クラスタ数が一定数より少なくなればク ラスタリングを終了し、そうでなければ 手順2へ戻る。

手順 2 の類似度の計算には、以下に示すような式が用いられます。

最近隣法

$$d_{\min}(x_p, x_q) = \min \left\| X - X' \right\|$$

最遠隣法

 $d_{\max}(x_p, x_q) = \max \left\| X - X' \right\|$ 

#### 重心法

$$d_{cnt}(x_p, x_q) = \left\| C_p - C_q \right\|$$

平均距離法

$$d_{avg}(x_p, x_q) = \sum \sum \frac{\|X - X'\|}{n_p n_q}$$

ここで、 $x_p$ 、 $x_q$  はクラスタ、X、X' はクラ スタ内の要素、 $C_p$ 、 $C_q$  はクラスタの重心、  $n_p$ 、 $n_q$ はクラスタ内の要素数を示します。





クラスタリングの結果は、使用する類似度計 算法やクラスタ数によって異なったものになり ます。最適な類似度計算法やクラスタ数は、対 象とするデータの特徴などにより変わるため、 類似度計算法やクラスタ数を変化させながら最 適な設定を求める必要があります。



第5章 クラスタリング





# 第6章 Back Propagation ニューラルネットワーク

この章では、代表的なニューラルネットワーク手法である、Back Propagation ニューラルネットワーク (BPNN) について説明します。 またChemishで BPNN によるモデリングを行う方法について説明し ます。

この章の内容

● BPニューラルネットワークとは
 ● BPニューラルネットワークの学習
 ● 学習回数-Q<sup>2</sup>プロットを表示する
 ● 中間層の最適ユニット数を求める
 ● 予測を行う

### **BPNNの概要**

Back Propagationニューラルネットワーク (BPNN) は、階層型のネットワーク構造をもっ た教師ありのニューラルネットワークです。最 急降下法を用いて実際の出力と希望の出力(教 師信号)の誤差が最小になるように結合荷重を 変更することで学習を行います。誤差をもとに 結合荷重の変更量を計算する様子から、誤差逆 伝播法と呼ばれます。

BPNN では中間層に非線形ユニット、出力 層に線形ユニットを用いた3階層型ネットワー クモデルを用いてモデリングを行います。出力 層のユニットを線形にすることにより、出力値 は-∞から+∞の値をとることが可能となり、 非線形モデリングを行うことができます。

下図は、BPNN の概念図です。



ネットワークは入力x<sub>1</sub>, x<sub>2</sub>, ・・・, x<sub>d</sub>に対して、 N個の中間層を経て出力層から値を出力します。



入力層→出力層の計算

次式によって、中間層のa番目のユニットでの総和S<sub>a</sub>を算出します。

$$S_a = \sum_{k=1}^d w_{ka} x_k + \theta_a$$

ここで、 $W_{ka}$ はk番目の入力層ユニットとa番目 の中間層ユニットとの間の結合荷重(重み)、 $\theta_a$ はa番目の中間層ユニットのバイアスです。

次にS<sub>a</sub>に対してシグモイド変換を行い、a番目の中間層ユニットは次式のt<sub>a</sub>を出力します。

$$t_a = \frac{1}{1 + \exp\left(-s_a / T\right)}$$

中間層の出力t<sub>a</sub>を出力層への入力とし、出力 層のI番目のユニットにおける総和O<sub>I</sub>を計算し ます。

$$O_l = \sum_{a=1}^A v_{al} t_a + \eta_l$$

ここで、 $\nu_a$ はa番目の中間層ユニットとl番目の出力層ユニット間の重みで、 $\eta_i$ はl番目の出力層ユニットのバイアスです。



# BPNNとは

### 結合荷重の修正

ネットワークの出力が求まったら、その出力 と教師信号との誤差関数Eが最小になるように 結合荷重とバイアスの修正を次式により行いま す。

$$E = \frac{1}{2} \sum_{l=1}^{m} (z_l - o_l)$$
  

$$\Delta v_{al} = \alpha (z_l - o_l) t_a$$
  

$$\Delta \eta_{al} = \beta (z_l - o_l)$$
  

$$\Delta w_{ka} = \alpha \left[ \sum_{l=1}^{m} (z_l - o_l) v_{al} \right] t_a (1 - t_a) x_k / T$$
  

$$\Delta \theta_{ka} = \beta \left[ \sum_{l=1}^{m} (z_l - o_l) v_{al} \right] t_a (1 - t_a) / T$$

ここで、z は教師信号、 $\Delta \nu_a$ は出力層への重 みの修正量、 $\Delta \eta_i$ は出力層のバイアスの修正 量、 $\alpha$ および $\beta$ は学習定数を表します。

以上の手順により、出力層ユニットから中間 層ユニットを経て入力層に向かって出力誤差 (z-o) が伝わって、重みとバイアスの修正が行 われます。この修正を誤差Eが十分小さくなる まで繰り返すことによって、ネットワークの学 習が行われます。



BPNNの学習を成功させるためには、いく つかのパラメータを適切に設定することが必 要になります。Chemish では、中間層のユ ニット数、学習回数、学習率などの値を設定 することが可能です。一般に誤差を減少させ るためには、中間層ユニットを多く、学習回 数を多く、学習率を高く設定します。

### 🦉 モデル評価

BPNNの学習が適切に行われたかどうかを 判断することは、それほど簡単ではありませ ん。中間層のユニット数、学習回数、学習率 などの値を高く設定することによって、学習 誤差の値を小さくすることはできますが、誤 差が小さいだけでは良いモデルであるとは断 言できません。学習に用いるデータに過剰に 依存してしまう、いわゆる過学習が起こるた めです。過学習状態のネットワークでは、外 部データについて精度の高い予測を行うこと ができません。

過学習を起こしていないかを確認するため には、学習に用いていない外部データを用い た予測テストや、Leave-one-outなどのクロ スバリデーションが有効です。またChemish には自動的に最適な中間層ユニット数を求め る機能も実装されています。













### h

# 学習回数-Q<sup>2</sup>プロットを表示する





1

学習の設定を行う

① [中間層の最適化] に チェックを入れる
BP計算条件設定
5 ▼ 中間層 重みの初期値を読み込む
100 文学習回数 参照
10 🔽 誤差計算間隔 🔽 クロスパリデーション
0.01 💌 学習率 🔽 重みの初期値を固定
0.80 モーメント係数 ア 中間層の最適化
120 マネットワーク温度 計算開始 キャンセル
② [計算開始] をクリック

### <mark>重みの</mark>初期値の固定

BPNNでは通常、学習前に重みをランダム に初期化しますが、Chemishでは重みの初期 値をファイルから読み込むことにより、ある 一定の値に設定することもできます。

重みの初期値のファイルを読み込むには、 まず重みのシートW1、W2があるCHSファイ ルを用意します。次にBPの計算条件設定ダ イアログの「重みの初期値を読み込む」とい うテキストの下にあるテキストボックスに CHSファイルのパスを入力します。また参照 ボタンをクリックすると、ファイルを選択す ることもできます。残りの計算条件の設定方 法は、普通に学習させるときと同様に行いま す。







Back Propagation 学習を行う場合には、 どの程度中間層のユニットを用意するかとい うことが重要になります。一般に中間層のユ ニットを多くすると教師データとニューラル ネットワークの出力の誤差は小さくなります が、ユニットの数が多すぎると学習データ (教師データ)の場合のみ出力の誤差が小さく なるオーバーフィッテイング(過学習)と呼ば れる状態になります。この問題を解決するた めに、Chemishでは自動的に最適な中間層ユ ニット数を求める機能を実装しています。





予測を行う



3 予測値が表示される

<ul> <li>***</li> <li>***</li> <li>**</li> <li>**</li> <li>**</li> <li>**</li> <li>*</li> <l< th=""><th>a 🛛</th><th></th><th></th><th></th><th></th><th>×</th></l<></ul>	a 🛛					×
変数値	説明変数	H(1)	H(2)	H(3)	H(4)	
80.000	前時	-0.819	0.604	0.620	-0.323	
10.000	テレビ	0.546	0.241	0.486	0.391	
40.000	睡眠	0.348	0.381	0.615	0.265	
80.000	前回の点数	-0.595	1.344	1.144	-0.008	
4.000	授業態度	0.309	0.321	0.800	0.274	
	バイアス	0.106	0.399	0.493	0.580	
予測値	_ 目的変数	H(1)	H(2)	H(3)	H(4)	バイアス
103.087	人。 使人	-1.466	1.141	0.526	-0.399	0.083

一括予測

Chemishでは学習済みのBPNNを用いて、 複数の入力データを用いた一括予測を行うこ とができます。一括予測を行うためには、ま ず予測させたい説明変数をCSVファイルとし て用意します。このときCSVファイルと学習 済みのBPNNの説明変数の数が同じになるよ うにします。

次にメニューから [計算]-[一括予測] をク リックすると、CSVファイルを開くためのダ イアログが表示されます。このダイアログで 予測させたい説明変数が格納されたCSVファ イルを開くと、一括予測を行った計算結果の シートが追加されます。





# 第7章 品質管理

この章では、Chemishで簡単な品質管理を行う方法について説明 します。

この章の内容

品質管理とは
 ヒストグラムの表示
 パレート図の表示
 管理図の表示

# 品質管理とは

### 品質管理の基礎

製品を大量に製造する場合、さまざまな要因 によってその品質にはばらつきが生じます。品 質管理とは、統計的な手法によりこのようなば らつきを解析し、消費者の満足する製品を一定 の品質を保ちつつ、なるべく低いコストで製造 するための手法です。

ー般に品質管理には、「QC 七つ道具」と呼ばれる以下の手法がよく用いられます。

- ・パレート図
- 特性要因図
- ・ヒストグラム
- 管理図
- ・チェックシート
- 散布図
- ・層別

Chemish ではこれらのうち、ヒストグラム、 パレート図、管理図、散布図を用いた分析を行 うことができます。散布図については、既に説 明が行われているので、そちらを参照してくだ さい。



- ・マトリクス図法
- 系統図法
- 連関図法
- 親和図法
- ・PDPC 法
- ・アローダイアグラム
- ・マトリクスデータ解析法

### 品質管理に関する文献

品質管理には、統計的手法の知識の他に、それらを使って得られた結果を考察するノウハウ が必要になります。このように品質管理には広 範な知識が必要になるため、興味のある方は初 めに以下に挙げたような入門書を参考にされる ことをお勧めします。

- 2. 奥村士郎、改訂版 品質管理入門テキスト、 日本規格協会 (1996)
- 石原勝吉 他、やさしいQC七つ道具、日本 規格協会(1980)







パレート図の表示





# 第8章 Kohonen ネットワーク、 Counter-Propagation

この章では、Kohonen の自己組織化ネットワークと、その拡張で ある Counter-Propagation について説明します。また、Chemishで Kohonen ネットワークや Counter-Propagation の学習を行う方法 について説明します。

この章の内容

Kohonenネットワークとは
 Kohonenネットワークの学習
 重みベクトルを表示する
 学習データを発火させる
 U-Matrixを見る
 CPとは
 CPの学習
 重みベクトルを表示する
 応答曲面を表示する



### Kohonenネットワークとは

#### 概要

Kohonenの自己組織化ニューラルネットワーク(Kohonen ネットワーク)は、Teuvo Kohonenによって1982年に発表された教師な し学習を行うニューラルネットワークの一種で す。Kohonen ネットワークは大脳視覚野での 情報処理を模倣したものであるといわれ、デー タ分類や情報圧縮に有効な手法として注目され ています。化学の分野においても、Kohonen ネットワークを用いた研究が数多く発表されて います。

Kohonen ネットワークでは下図に示すよう な 2 次元のマップを用いて学習を行います。 この図では、それぞれの円がニューロンを示し ており、各ニューロンは重みベクトルを持って います。この図では重みベクトルを矢印で表現 しています。学習の目的は図右に示すような 2 次元マップを得ることであり、これにより多次 元情報のクラスタリングや次元縮約を行うこと ができます。右にKohonenネットワークの学習 アルゴリズムを示します。



### 学習アルゴリズム

- 1. マップ上のすべての重みベクトルをランダム に初期化する
- 2. 参照する入力ベクトルを一つ選択する
- マップ上の全ニューロンの中から、入力ベクトルに最も類似する重みベクトルを持ったニューロン(勝者ニューロン)を選択する。 一般には類似度としてユークリッド距離を用いる
- 勝者ニューロンとその近傍ニューロンの重み ベクトルを、入力ベクトルに近づくように修 正する
- 5. 設定した学習回数に達していれば終了する。 そうでなければ手順2へ戻る

以上の手順により、Kohonen ネットワークは 入力空間でのサンプルデータの分布を学習し、入 力空間での相対的な位置関係(トポロジー)を保 持したままサンプルデータを二次元平面上のマッ プに写像します。Kohonen ネットワークにおい ても、他のニューラルネットワークと同様に、学 習に用いるパラメータを適切に設定することが必 要になります。主なパラメータとしては、学習回 数、マップの大きさ、学習率などがあります。


















## 概要

Counter-Propagation手法は、Hecht-Nielsen が発表したKohonenネットワークの拡張モデル で、教師ありニューラルネットワークの一種で す。多次元データを用いた非線形モデリングを 行うことができます。学習の結果については、 Kohonenネットワークにおける重みベクトルの 等高線図と同様に、3次元グラフとして視覚化 することが可能です。

Counter-Propagationネットワークは、下図 のようにKohonenネットワークの2次元マップ にもう一枚出力層を加えた構造を持っています。 そのマップは応答平面と呼ばれ、Kohonenネッ トワークの2次元マップと同時に学習が行われ ます。以下にCounter-Propagationネットワー クの学習アルゴリズムを示します。 学習アルゴリズム

- Kohonenネットワークと同様に、入力ベク トルに最も類似する重みベクトルを持つ勝 者ニューロンを探索する
- 勝者ニューロンとその近傍のニューロンの 重みベクトルを入力ベクトルに近づく方向 へと修正する
- 応答平面上において勝者ニューロンと同じ 位置にあるニューロンとその近傍ニューロ ンの重みベクトルを、教師信号ベクトルに 近づく方向へと修正する
- 設定した学習回数に達していれば終了する。
  そうでなければ手順1へ戻る















## 第9章 逆解析

この章では、Chemishで逆解析を行う方法について説明します。

この章の内容

● 逆解析とは ● 候補を提案する ● 候補のスクリーニングを行う





第9章 逆解析

候補を提案する









計算間隔は、コンボボックスになっていま す。セルをダブルクリックすることによって、 1/5から1/100までの候補が表示されますので、 そのなかから計算間隔を指定します。例えば 1/5を指定すると、最小値から最大値の間が5 等分されて、その各値を用いた予測値計算が 行われます。また計算値は、計算間隔を値で 指定するためのものです。例えば、1.0を指 定すれば、1.0間隔で予測値計算が行われま す。



8

9

## 候補を提案する

計算開始





逆解析の計算結果です。[RApar] には逆解析 を行った際の計算条件が記述されています。 全探索を行う際の各説明変数の範囲、計算間 隔、計算値などです。[RA] には、逆解析に よって得られた候補が記述されています。 [RA] の右端の列は得られた候補がどの程度 条件を満たしているかを示す値"スコア"が 格納されています。スコアが10に近いほど条 件に一致した候補であるといえます。

**9** 結果を見る

逆解析の計算結果のシートが追加される

😋 Ch	emish	- [test.csv]							_ <b>_ _ _</b> ×
5 🖸	ァイル	編集表	示 シート	計算 QC	グラフ ツ	ール ウベ	パウ へル	1	_ 8 ×
D.	2 E		PIS RD	<b>N</b> + 18	I 🖌 I m	9			
		3 121141			-   baba	•			
XY	B	VAT	Ycalc	R2 Par	RA	RApar			
		勉強	テレビ	睡眠	前回の点数	授業態度	点数	スコア	<b>▲</b>
	1	70.000	0.000	49.400	47.200	1.000	97.674	9.767	
	2	70.000	10.400	49.400	47.200	1.000	97.673	9.767	
	3	70.000	20.800	49.400	47.200	1.000	97.672	9.767	
	4	70.000	31.200	49.400	47.200	1.000	97.671	9.767	
	5	70.000	41.600	49.400	47.200	1.000	97.670	9.767	
	6	70.000	52.000	49.400	47.200	1.000	97.669	9.767	
	7	70.000	0.000	35.000	64.800	1.000	97.086	9.709	
	8	70.000	10.400	35.000	64.800	1.000	97.085	9.708	
	9	70.000	20.800	35.000	64.800	1.000	97.084	9.708	
	10	70.000	31.200	35.000	64.800	1.000	97.083	9.708	
	11	70.000	41.600	35.000	64.800	1.000	97.082	9.708	
	12	70.000	52.000	35.000	64.800	1.000	97.081	9.708	
	13	56.000	0.000	49.400	64.800	1.000	96.494	9.649	
	14	56.000	10.400	49.400	64.800	1.000	96.493	9.649	
	15	56.000	20.800	49.400	64.800	1.000	96.492	9.649	
	16	56.000	31.200	49.400	64.800	1.000	96.491	9.649	
	17	56.000	41.600	49.400	64.800	1.000	96.491	9.649	
	18	56.000	52.000	49.400	64.800	1.000	96.490	9.649	
	19	56.000	0.000	35.000	82.400	1.000	95.906	9.591	
	20	56.000	10.400	35.000	82.400	1.000	95.905	9.591	
	21	56.000	20.800	35.000	82.400	1.000	95.905	9.590	-
1 X_7 9.767354									NUM

















🥊 マップの評価

逆解析により発生する候補は、計算条件に よっては非常に大量になる場合があります。 そこで、候補をKohonenマップに発火させ、 その発火位置によりスクリーニングおよび評 価を行います。例えばクラス3のY\_Wのグラ フ上にクラス3の候補を発火させた場合を考 えます。このとき、赤い部分に発火する候補 は、学習データにおける評価値の高いサンプ ルと似たような説明変数を持っていることが 分かります。つまり既存のものと大差が無く、 新規性はあまり期待できないと考えられます。 逆に青い部分に発火した候補は、学習サンプ ルの中に似たような説明変数を持ったものが なく、新規性が高く、かつ高い活性を示すこ とが期待できます。

