

College Analysis 総合マニュアル

－ O R 2 －

目次

| | |
|-----------------------|----|
| 8. システムダイナミクス | 1 |
| 9. 不良品診断（以降解答未） | 11 |
| 10. パラメータ設計 | 19 |
| 11. オンライン品質工学 | 31 |
| 12. 異常検知..... | 42 |
| 13. ナッシュ均衡ツール | 56 |

8. システムダイナミクス

8.1 システムダイナミクスとは

システムダイナミクスは 1950 年代から発展してきたシミュレーション手法で、有益性は広く認識されている。ソフトウェアも STELLA/iThink (以後 STELLA と呼ぶ) や Dynamo を代表として、多く提供されており、社会システム分析の一分野を形成している。元々この手法は、シミュレーション記述言語によって分析を行っていたが、コンピュータの発達によって、グラフィックによる記述が一般的になり、見た目にも分かり易い表現となった。我々は、このグラフィック表現をグラフィックエディタによって描画することを考え、College Analysis への導入を検討した。プログラム作成に当っては参考文献[1] の STELLA の描画を参考にしたが、これまでのグラフィックエディタの利用法との相違がないようにした。また、外界を表すクラウドについては、開発当初は導入していたが、グラフィック画面の簡単化のために省略している。

グラフィックエディタにおける画面の構成要素は、ボックスとラインである。そのため、フロー（バルブを含む）のように一体化した考え方はない。また、グラフィックエディタでは、ボックス間を結ぶラインは、方向が同じものは 1 本だけしか定義されない。この点が STELLA と最も大きな違いであるが、これはライン間に節点と呼ばれるボックスを 1 つ挟むことで解決した。この節点は、KJ 法の描画の際に導入したもので、コピーや印刷の際には表示されないの、うまく工夫すると描画から消し去ることもできる。

システムダイナミクスを取り扱う際の、グラフィックエディタでのボックスは、ストック、バルブ、コンバータであり、ラインは、フロー（直線矢印）とコントロール（左右円弧矢印）である。ボックスにおけるデータは、数値や数式の形で、ボックスの value の中に保存される。数式の表示では、ファイルデータや n 期前のデータなどの参照が可能である。

8.2 プログラムの利用法

メニュー [分析－OR－システムダイナミクス] を選択すると、図 1 のようなシステムダイナミクス実行画面が表示される。

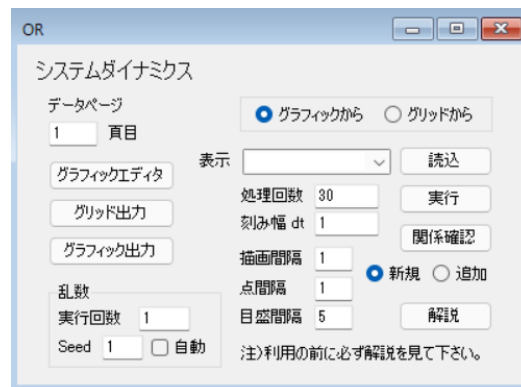


図 1 システムダイナミクス実行画面

最初にこの画面で「グラフィックエディタ」ボタンをクリックして、図 2 のグラフィックエディタ画面を表示する。

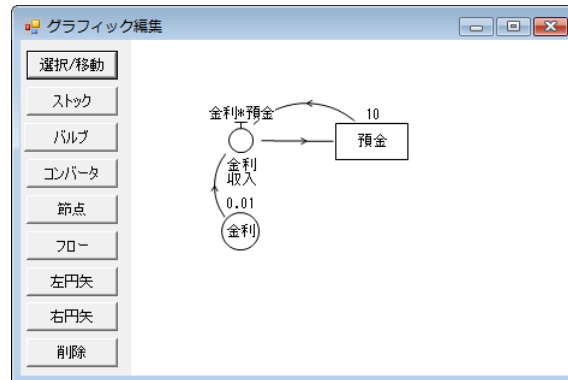


図 2 グラフィックエディタ画面

描かれた図は金利収入による預金の増加モデルである¹⁾。預金は「ストック」と呼ばれる貯蔵庫、金利は「コンバータ」と呼ばれる影響を表す記号、金利収入は「バルブ」と呼ばれる制御機能を表す記号で表示されている。金利収入から預金への直線は「フロー」と呼ばれ、ものの流れ（この場合は金銭の流れ）を表している。バルブはこのフローを制御している。金利から金利収入、または預金から金利収入への円弧は「コントロール」と呼ばれ、コンバータやストックから金利収入への影響を表している。図の中では、金利から金利収入への円弧は「左円矢」、預金から金利収入へは「右円矢」を利用している。二つの円形のラインは進行方向左側の円弧か、右側の円弧かの違いだけで同じものであり、見た目にバランスの取れた方を利用すればよい。

ボックス名の入力にはボックスをダブルクリックし、数値や数式を表す **value** の設定は右クリックで入力する。図 2 では、金利の **value** は「0.01」、金利収入の **value** は「金利*預金」、預金の **value** は初期値の 10（万円）に設定している。

グラフィックエディタメニュー [表示-ボックス値表示[ON/OFF]] で表示のモードを切り替えると、図 3 のようにボックスの上の **value** 値を非表示にすることができる。

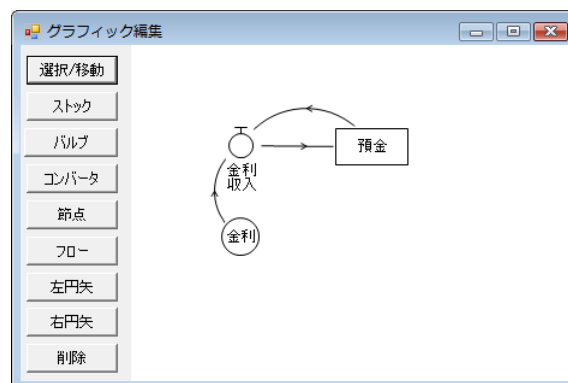


図 3 ボックス値非表示モード

できたデータは、「グリッド出力」ボタンをクリックしてグリッドエディタに保存することができる。分析の実行は、「グラフィックから」、「グリッドから」ラジオボタンで、グラ

フィックデータを直接利用するか、保存したグリッドデータを利用するか選択する。通常は「グラフィックから」にしておき、必要に応じてグリッドエディタに保存する。新規の場合やデータの構造を変えた場合は、「読込」ボタンでプログラム中にデータを読み込み、「表示」コンボボックスで、出力するデータを選択し、「実行」ボタンをクリックする。データの構造を変えず、**value** のみ変更した場合は、そのまま「実行」ボタンをクリックしてもよい。「表示」コンボボックスで預金を選んだ実行結果を図 4 に示す。

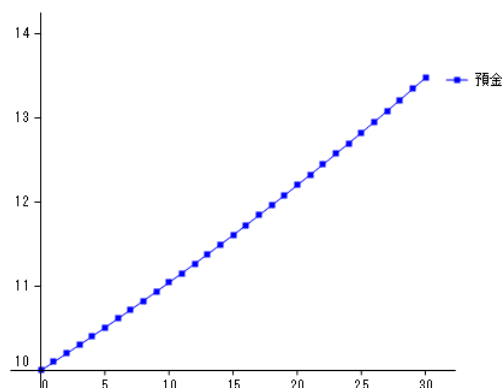


図 4 預金表示結果

数式や数値の評価で、コンバータとバルブはその時点で計算を行うが、ストックに関しては、常に 1 期前のデータを参照している。

金利を、0.01, 0.02, 0.03 と変えた場合の比較は、**value** を変更しながら、実行画面下の「追加」ラジオボタンを選択してグラフを描画する。その結果を図 5 に示す。ここで凡例は変更してある。

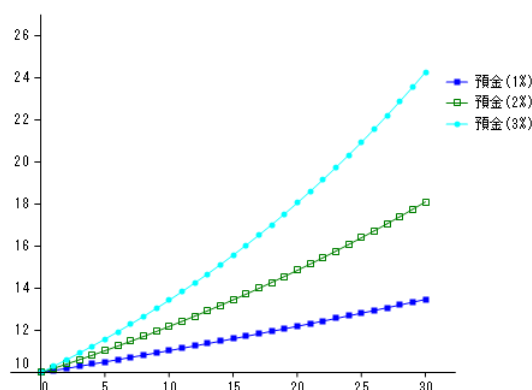


図 5 金利を変化させた試行

「関係確認」ボタンをクリックすると、図 6 のようなボックス間の関係がまとめて表示される。資料などに使用する際に利用する。

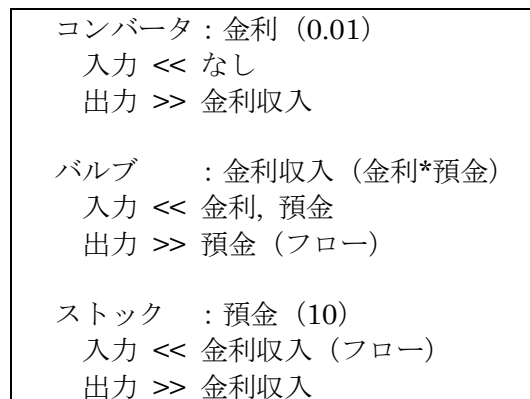


図 6 ボックス間の関係

次にアルコールの吸収・排出問題の例である^[1]。図 7a にグラフィックエディタの画面、図 7b にそれをコピーしたものを示す。

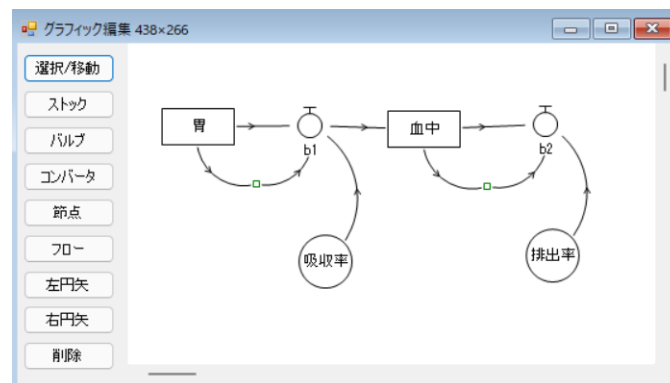


図 7a グラフィックエディタ画面

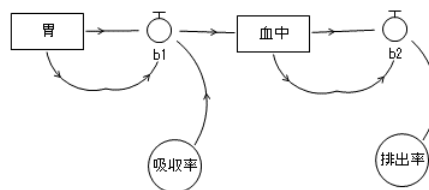


図 7b グラフィックデータのコピー

これは胃からバルブ b1 に、血中からバルブ b2 に、フローとコントロールの 2 本のラインが入る例であるが、グラフィックエディタの仕様から、同じボックス間に同方向の複数のラインを引くことができず、間にグラフィックエディタの左側にボタンのある「節点」を入れている。節点はコピー画像では消えている。しかし、この節点が気になる場合は、図 8a のようにボックスの内部に節点を持って行く。コピーすると図 8b のようにあたかも節点が無かったかのように見える。

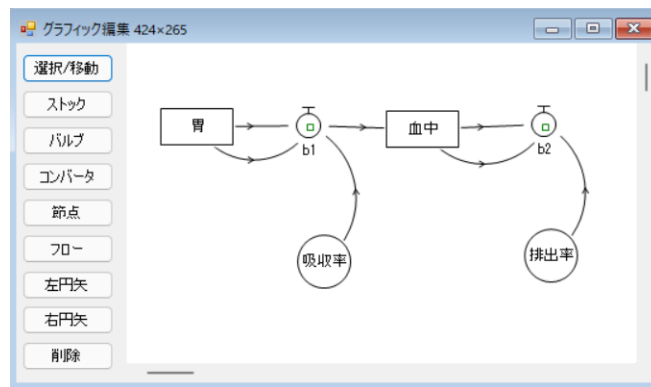


図 8a グラフィックエディタ画面

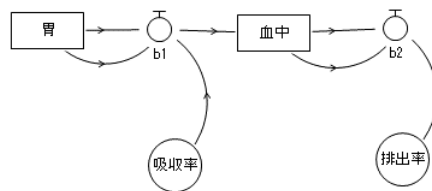


図 8b グラフィックデータのコピー

図 9 に value 値を表示したグラフィックデータ、図 10 に「実行時間」を 100、「目盛間隔」を 10 とした実行結果を示す（見易いように、排出率の値は参考文献と変えてある）。節点があっても、それについての設定は不要である。

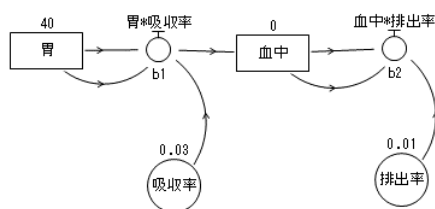


図 9 value 値付きグラフィックデータ

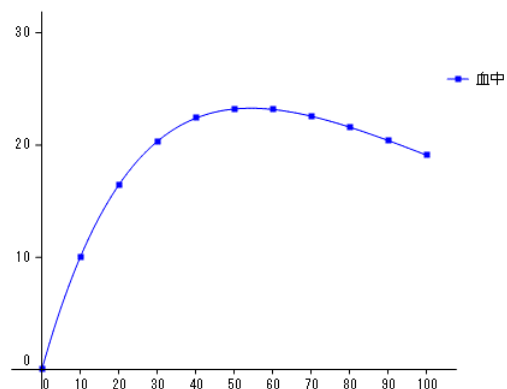


図 10 実行結果

次にファイルのデータを利用する在庫管理問題の例を考える。図 11 がそのモデル画面である。ファイルのデータは出庫データで、このモデルのデータが入っているファイルの 1 ページ目に図 12 のように含まれている。

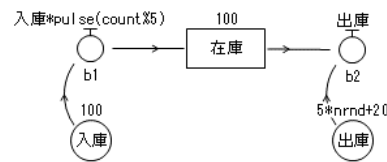


図 11 ファイルデータを利用した在庫の変化 (p2)

| | 出庫1 | 出庫2 | 出庫3 |
|---|-----|-----|-----|
| 1 | 18 | 34 | 10 |
| 2 | 17 | 25 | 13 |
| 3 | 14 | 30 | 14 |
| 4 | 17 | 24 | 12 |
| 5 | 19 | 41 | 16 |

図 12 出庫データ

入庫量は常に 100、入庫は 5 日ごとにする。バルブ b1 の式「 $\text{入庫} * \text{pulse}(\text{count} \% 5)$ 」の count 変数は経過回数で 1 から始まる。「 $\text{count} \% 5$ 」は count を 5 で割った余りを表す。pulse(x) 関数は、 $x=0$ のとき 1、その他のとき 0 の値を持つ。これを文章で表すと、「時間を 5 で割った余りが 0 のときだけ、入庫(100)で表わされる値を在庫ストックに送り込む」ということになる。出庫コンバータの式「 $\$ \text{出庫} 1$ 」で、変数名の先頭に付いた \$ マークは、分析メニューの「データページ」テキストボックスで与えたページにあるデータを参照することを意味する。変数名は「出庫 1」であり、経過時間 t に t 行目のデータを参照する。在庫量のシミュレーション結果を図 13 に示す。

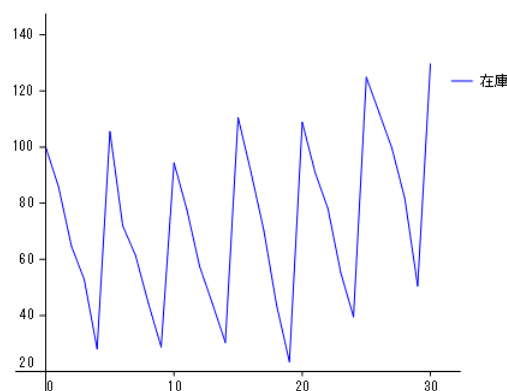


図 13 在庫量シミュレーション結果

グラフ画面のメニュー「設定—データポイント[ON/OFF]」で、データ点は消してある。

次は平均 20、標準偏差 5 の正規出庫の場合の、定期発注方式のシミュレーションである。
図 14 にモデルを示す。

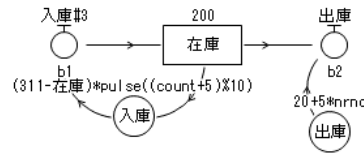


図 14 乱数出庫の定期発注方式 (p5)

出庫は「 $20+5*nrnd$ 」で表わされるが、「nrnd」は標準正規乱数であるので、出庫は平均 20、標準偏差 5 の正規乱数データとなる。入庫は 10 日ごとに発生するように pulse() 関数を用いて設定されているが、その大きさは、最大在庫 311 から現在の在庫の量を引いたものである。但し、在庫納入はバルブ b1 によって、入庫発生から 3 日後となる。b1 の式「入庫#3」は 3 日前の入庫データを参照することを意味する。ストックはその 1 日前の値を用いるので、この式の在庫は 4 日前の値である。そのため、入庫のリードタイムは 4 日と解釈される。在庫量のシミュレーション結果を図 15 に示す。

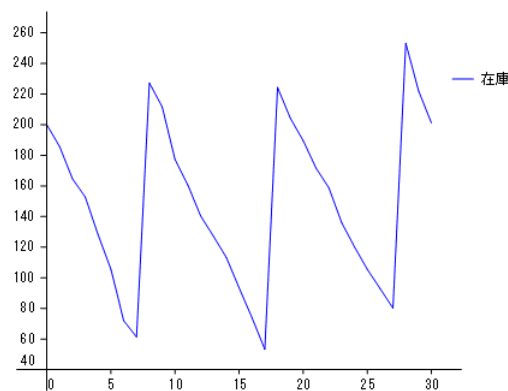


図 15 在庫量シミュレーション結果

乱数は、図 1 のメニューにあるように、Seed を決めて、同じ乱数系列を発生させることも、自動的に Seed を決めて毎回違う乱数系列にすることもできる。メニューの実行回数は、今後拡張するモンテカルロシミュレーション用に作った部分であるが、現在はまだ使用していない。

定量発注方式の表現は、定期発注方式より少し複雑である。在庫が発注点在庫 (96) より少なくなった時点で発注するので、図 16 の入庫コンバータのような複雑な式になる。theta(x) 関数は $x \geq 0$ のとき 1、 $x < 0$ のとき 0 となる関数である。pulse() 関数と合わせて使うと、条件文などが表現できるようになる。図 17 に定量発注方式の在庫量シミュレーション結果を示す。

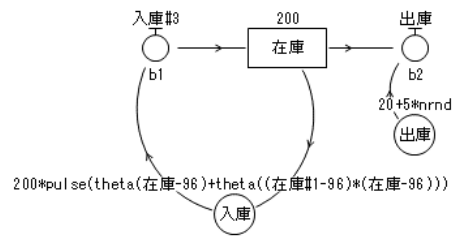


図 16 乱数出庫の定量発注方式 (p4)

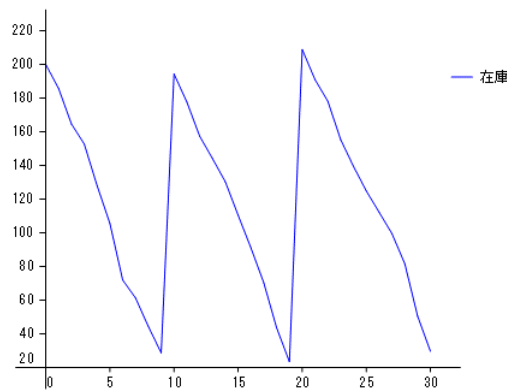


図 17 在庫量シミュレーション結果

これまでではストックに上限や下限はなかったが、新しくこれを設定できるようにした。例えば、ストックの値として、100;0;200のように、セミコロンで区切ると、初期値 100、下限 0、上限 200 となる。また、100;0; や 100;;200 のように、上限や下限を略すとその部分の制約は付かなくなる。この制約は現在のバージョンでは、単独のバルブを介する場合のストック間に有効で、1つのストックに複数のバルブが繋がった場合の制御はまだ完全ではない。図 18 と図 19 に最も簡単な制約付きストックの例を示す。

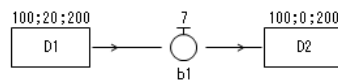


図 18 制約付きストック

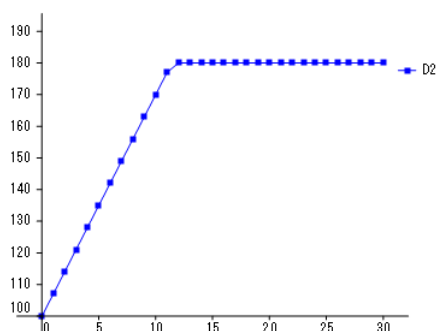


図 19 制約付きストックシミュレーション結果

次に常微分方程式の解法についてプログラムを書き直した。以下の常微分方程式について、システムダイナミクスのプログラムで解くことを考える。

$$dy/dt = -y + \sin(t)$$

初期条件： $t=0$ において $y=5$

システムダイナミクスのダイアグラムは通常図 20 で表す。

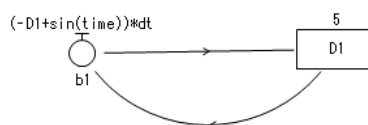


図 20 微分方程式（オイラー法）

ここで、「dt」は分析メニューに含まれる「刻み幅」で、利用者が設定する。また、「time」はそれまでの経過時間である。この方法は、オイラー法と呼ばれる解法である。結果を、図 21 に示す。

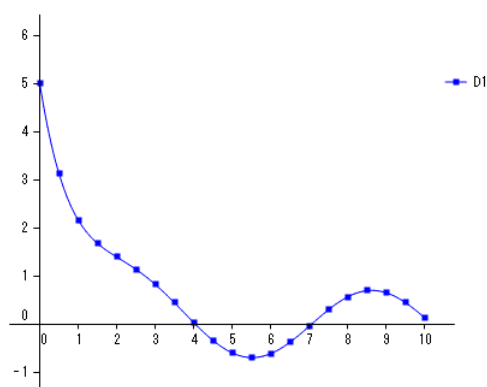


図 21 微分方程式の解（オイラー法）

ここでは、実行メニューの「処理回数」1000、「刻み幅」0.01、「描画間隔」10、「点間隔」5、「目盛間隔」10 に設定してある。ここで、点間隔と目盛間隔は、処理回数と描画間隔で

与えた描画点の数 $1000 \div 10 = 100$ の中での間隔である。

さて、通常の常微分方程式の解法では、ルンゲ・クッタ法がよく利用される。C. Analysis の常微分方程式のプログラムでもルンゲ・クッタ法を利用している。そのため、システムダイナミクスでもルンゲ・クッタ法を用いることを考えてみた。それが図 22 の方法である。

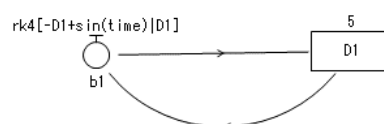


図 22 微分方程式の解（ルンゲ・クッタ法）

ここで、 $rk4[-D1+\sin(\text{time})|D1]$ が、図 20 の $(-D1+\sin(\text{time})) \cdot dt$ の代わりである。どの変数が微分されているかをはっきりさせるために、「 $|D1$ 」の部分が必要である。「 $|$ 」より左に微分されている変数が含まれない場合は、特に加える必要はない。実際動かしてみると、文字列処理が増えるため多少計算時間がかかるが、1000 回の処理で 2 秒弱であった。もちろん通常のルンゲ・クッタ法の常微分方程式のプログラムに比べて時間はかかるが、複雑な処理もできる可能性があるため、少し期待できるのではないかと思う。

参考文献

- [1] シミュレーションによるシステムダイナミクス入門，土金達男，東京電機大学出版局，2005.

9. 不良品診断（以降解答未）

ここでは不良品の発生に関する2つのプログラムを紹介する。1つは不良率診断についてのプログラムで、2項分布を使った正確な適合度検定を使い易くしたプログラムである。これは過去の実績から、新しい不良品の発生頻度の確率を求めるものである。もう一つは、抜き取り検査についてのプログラムである。抜き取り検査は多数のロットの中から、サンプルを取り出し、その中の不良品数でロットの合否を決めるものである。プログラムは、サンプル数、合格判定数、ロット中の不良品率、ロットの合格率のうちから、3つのデータを指定し、残りの1つを計算する機能とOC曲線を描く機能を持つ。

9.1 不良率診断

不良品の発生確率を過去の実績から求めることは生産ラインの異常を知る上で重要である。発生確率が小さい場合、発生は偶然とは考えられない。

例えばこれまで不良品の発生確率が p であったとする。これが今回の観測では N 個中 n 個であったとする。この事象の起こる確率 P は、通常の検定では自由度1の χ^2 分布の性質を使った以下の式を利用する。

$$\chi^2_1(P) = \frac{(n - pN)^2}{pN} + \frac{(n - pN)^2}{(1 - p)N} \quad (1)$$

しかし、これは p と n が、ある程度大きい場合の近似式で不良品の発生などの少ない発生件数では使えない。正確な計算を行うには以下の2項分布の式を利用する。

$$P = \sum_{i=n}^N {}_N C_i p^i (1 - p)^{N-i} \quad (2)$$

プログラムでは、(2)式を使って計算している。

メニュー「分析－OR－品質管理－不良品診断」を選択すると図1に示す分析実行画面が表示される。データの入力方法は「ファイルから一括処理」と「入力データから」を選択でき、データの種類の「過去発生数から」と「過去発生率から」を選択できる。図1の左がファイルからで、右が入力データからの処理である。

図1 不良率診断の分析実行画面

例えば右の図のように、「データ」テキストボックスに、データを入力し、「分析実行」ボタンをクリックすると図 2 のような結果が表示される。

| | 過去売上数 | 過去不良数 | 過去不良率 | 売上数 | 不良数 | 不良率 | 検定確率 |
|---|-------|-------|--------|-----|-----|--------|----------|
| ▶ | 1000 | 4 | 0.0040 | 100 | 2 | 0.0200 | 0.061226 |

図 2 出力結果

データファイルの形式は図 3 のように、過去の個数、過去の不良品数、今回の個数、今回の不良品数を商品ごとに並べて入力する。過去の不良品率の場合は、過去の個数と不良品数の代わりにそれだけがあればよい。

| | 過去個数 | 過去Claim | 今年個数 | 今年Claim |
|-----|------|---------|------|---------|
| a | 1000 | 5 | 200 | 4 |
| b | 2000 | 7 | 300 | 5 |
| ▶ c | 1000 | 4 | 135 | 2 |

図 3 データファイル形式

分析実行メニューで図 4 のように変数を選んで、「ファイルから一括処理」ラジオボタンを選択し、「分析実行」ボタンをクリックすると、図 5 のような結果が表示される。

図 4 分析実行画面（ファイルの変数選択）

| | 過去売上数 | 過去不良数 | 過去不良率 | 売上数 | 不良数 | 不良率 | 検定確率 |
|-----|-------|-------|--------|-----|-----|--------|----------|
| ▶ a | 1000 | 5 | 0.0050 | 200 | 4 | 0.0200 | 0.018679 |
| b | 2000 | 7 | 0.0035 | 300 | 5 | 0.0167 | 0.004393 |
| c | 1000 | 4 | 0.0040 | 135 | 2 | 0.0148 | 0.102275 |

図 5 ファイルから一括処理の出力結果

9.2 抜き取り検査

抜き取り検査は、大きさ N 個のロットの中から n 個のサンプルを取り出して、その中に

含まれる不良品の数が c 個以下ならそのロットは合格、それより多い場合はそのロットを不合格にする品質検査の手法である。今このロットの中に確率 p で不良品が含まれているとき、サンプルの中から不良品が x 個見つかる確率 $p(x)$ は以下の超幾何分布の確率で与えられる。

$$p(x) = \frac{{}^pN C_x \cdot (1-p)^N C_{n-x}}{{}^N C_n}$$

特に $N \gg n$ で、ロットの中から取り出す個数の減少の影響を考慮しなくても良い場合、確率 $p(x)$ は以下の 2 項分布の確率で与えられる。

$$P(x) = {}_n C_x p^x (1-p)^{n-x}$$

これより、ロットが合格する確率 P は以下となる。

$$P = \sum_{x=0}^c p(x)$$

メニュー「分析－OR－品質管理－抜き取り検査」を選択すると、図 6 のような分析実行画面が表示される。ここで「ロットサイズ利用」チェックボックスを選択すると、超幾何分布を用いて計算する。

品質管理

抜き取り検査

☐ ロットサイズ(N)利用 1000

不良品率(p) 0.05 Clear

サンプル数(n) 20 Clear

不良品許容数(d) 1 Clear

ロット合格率(Q) 0.9 Clear

空欄を他のところから計算します。 計算

x 軸 不良品率

y 軸 ロット合格率

グラフ

x,y軸以外は上の値で固定して描きます。
カンマ区切りで複数本のグラフを描けます。

注)不良品許容数は合格判定個数のことです。
不良品率は不適合率のことです。

図 6 抜き取り検査の分析実行画面

上の 4 つのテキストボックスは、そのうちのどれかを空欄にして、その下にある「計算」ボタンをクリックすると計算結果で空欄を埋めてくれる。例えばこの状態で、「ロット合格率」を空欄にして、「計算」ボタンをクリックすると、計算結果の 0.7358 でテキストボックスが埋まる。また、「ロット合格率」を 0.9 にして、「サンプル数」を空欄にして「計算」ボタンをクリックすると、「サンプル数」で 11 という答えが表示される。この結果は図 7 に示す。

図7 3つの条件で他の条件を決める方法

メニューの条件で、不良品率を変化させた場合のロット合格率の変化を見たいときは、図7の設定で「グラフ」ボタンをクリックすると、図8のような結果が表示される。

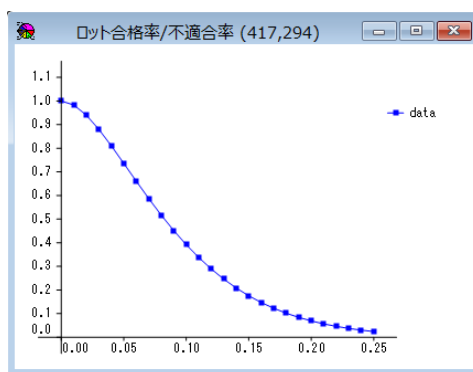


図8 不良品率の変化によるロット合格率の変化

さらに不良品許容数を 0,1,2,3 と変化させたいときは、「不良品許容数」にカンマ区切りでこのまま入力し、「グラフ」ボタンをクリックする。結果は図9のように表示される。

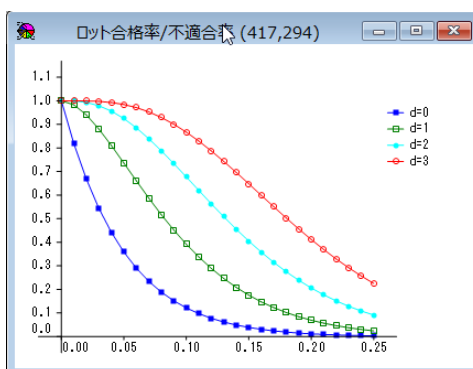


図9 不良品許容数を変化させたロット合格率の変化

同じ設定で、ロット合格率と不良率を逆にすると図10のような結果になる。

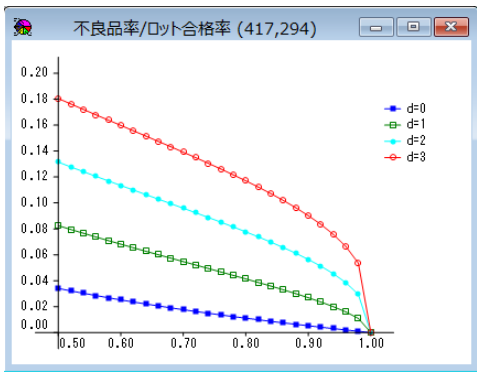


図 10 ロット合格率の変化による不良率の変化

問題 1（不良品診断）

不良品発生頻度に以下の表データ（品質管理（不良率診断）.txt）が与えられた。

| | 過去個数 | 過去 Claim | 今年個数 | 今年 Claim |
|---|------|----------|------|----------|
| A | 1000 | 5 | 200 | 4 |
| B | 2000 | 7 | 300 | 5 |
| C | 3000 | 8 | 500 | 2 |

このデータを用いて以下の問いに答えよ。

1) 今年の発生確率とその起きる確率である検定確率を求めよ。

| | 今年発生率 | 検定確率 |
|---|-------|------|
| A | | |
| B | | |
| C | | |

2) 有意水準を 5%と 1%としたときどのように判断するか。

| | 5% | 1% |
|---|-------|-------|
| A | 正常・異常 | 正常・異常 |
| B | 正常・異常 | 正常・異常 |
| C | 正常・異常 | 正常・異常 |

問題 2（抜き取り検査）

ロットからの抜き取り検査で、ロット内の製品数 N 、ロット内製品の不良品率 p 、取り出すサンプル数 n 、不良品許容数 d 、その際のロット合格率 Q がそれぞれ以下のように与えられるとき、問いに答えよ。

ロット内の製品数が十分多いとする場合（2 項分布）

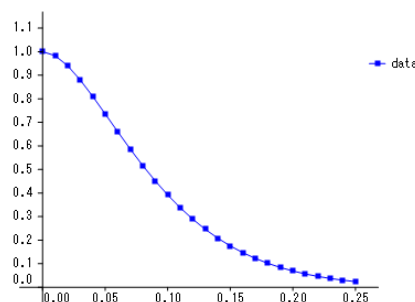
- 1) 不良品率 0.05、サンプル数 20、不良品許容数 1 のとき、ロット合格率はいくらになるか。ロット合格率 []
- 2) ロット合格率 0.9、サンプル数 20、不良品許容数 1 のとき、ロット内製品の不良品率はいくらか。不良品率 []
- 3) ロット合格率 0.9、サンプル数 20 で、不良品率を 0.05 以下にしたいとき、不良品許容数をいくつ以下にすればよいか。
不良品許容数を [] 以下にする。
- 4) ロット合格率 0.9、不良品許容数 1 で、不良品率を 0.05 以下にしたいとき、サンプル数をいくつ以上にすればよいか。
サンプル数を [] 以上にする。

ロット内製品数を 100 にする場合（超幾何分布）

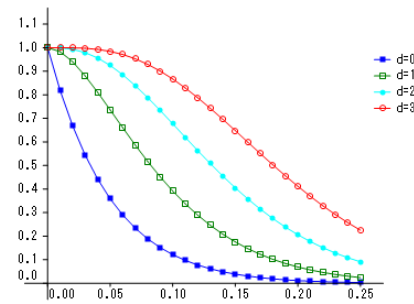
- 5) 不良品率 0.10、サンプル数 10、不良品許容数 1 のとき、ロット合格率はいくらになるか。ロット合格率 []
- 6) ロット合格率 0.95、サンプル数 10、不良品許容数 1 のとき、ロット内製品の不良品率はいくらか。不良品率 []
- 7) ロット合格率 0.95、サンプル数 10 で、不良品率を 0.1 以下にしたいとき、不良品許容数をいくつ以下にすればよいか。
不良品許容数を [] 以下にする。
- 8) ロット合格率 0.95、不良品許容数 1 で、不良品率を 0.1 以下にしたいとき、サンプル数をいくつ以上にすればよいか。
サンプル数を [] 以上にする。

注) 不良品許容数は通常合格判定個数と呼ばれます。不良品率は不適合率とも呼ばれます。もう一度、ロット内の製品数が十分多いとする場合（2 項分布）

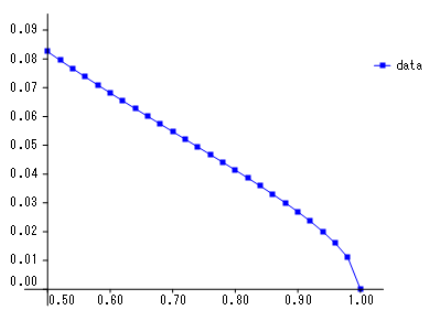
- 9) サンプル数 20、不良品許容数 1 として、不良品率を変化させたとき、ロット合格率はどのように変化するか。グラフで求めよ。



- 10) サンプル数 20、不良品許容数 0, 1, 2, 3 として、不良品率を変化させたとき、ロット合格率はどのように変化するか。グラフで求めよ。



11) サンプル数 20、不良品許容数 1 として、ロット合格率を変化させたとき、不良品率はどのように変化するか。グラフで求めよ。



問題 1 解答（不良品診断）

1) 今年の発生確率とその起きる確率である検定確率を求めよ。

| | 今年発生率 | 検定確率 |
|---|--------|--------|
| A | 0.0200 | 0.0187 |
| B | 0.0167 | 0.0044 |
| C | 0.0040 | 0.3851 |

2) 有意水準を 5%と 1%としたときどのように判断するか。

| | 5% | 1% |
|---|-------|-------|
| A | 正常・異常 | 正常・異常 |
| B | 正常・異常 | 正常・異常 |
| C | 正常・異常 | 正常・異常 |

問題 2（抜き取り検査）

ロット内の製品数が十分多いとする場合（2 項分布）

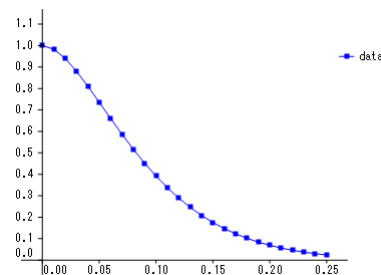
- 不良品率 0.05、サンプル数 20、不良品許容数 1 のとき、ロット合格率はいくらになるか。ロット合格率 [0.7358]
- ロット合格率 0.9、サンプル数 20、不良品許容数 1 のとき、ロット内製品の不良品率はいくらか。不良品率 [0.0269]
- ロット合格率 0.9、サンプル数 20 で、不良品率を 0.05 以下にしたいとき、不良品許容数をいくつ以下にすればよいか。
不良品許容数を [1] 以下にする。
- ロット合格率 0.9、不良品許容数 1 で、不良品率を 0.05 以下にしたいとき、サンプル数をいくつ以上にすればよいか。
サンプル数を [11] 以上にする。

ロット内製品数を 100 にする場合（超幾何分布）

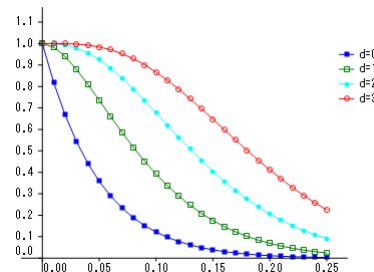
- 5) 不良品率 0.10、サンプル数 10、不良品許容数 1 のとき、ロット合格率はいくらになるか。ロット合格率 [0.7385]
- 6) ロット合格率 0.95、サンプル数 10、不良品許容数 1 のとき、ロット内製品の不良品率はいくらか。不良品率 [0.0450]
- 7) ロット合格率 0.95、サンプル数 10 で、不良品率を 0.1 以下にしたいとき、不良品許容数をいくつ以下にすればよいか。
不良品許容数を [2] 以下にする。
- 8) ロット合格率 0.95、不良品許容数 1 で、不良品率を 0.1 以下にしたいとき、サンプル数をいくつ以上にすればよいか。
サンプル数を [5] 以上にする。

もう一度、ロット内の製品数が十分多いとする場合（2 項分布）

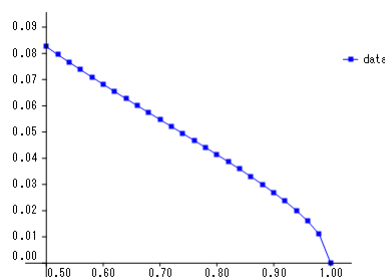
- 9) サンプル数 20、不良品許容数 1 として、不良品率を変化させたとき、ロット合格率はどのように変化するか。グラフで求めよ。



- 10) サンプル数 20、不良品許容数 0, 1, 2, 3 として、不良品率を変化させたとき、ロット合格率はどのように変化するか。グラフで求めよ。



- 11) サンプル数 20、不良品許容数 1 として、ロット合格率を変化させたとき、不良品率はどのように変化するか。グラフで求めよ。



10. パラメータ設計

10.1 パラメータ設計とは

装置からの出力は、人が制御できる制御因子と制御できない誤差因子に影響され、理想値からのずれが生じる。パラメータ設計とは、データの誤差因子の影響を抑え、理想値に近い観測値が得られるよう直交表と呼ばれる組み合わせ表を用いて制御因子を調整する手法である。パラメータ設計には以下に示す動特性と静特性がある。

動特性：入力に対して理想的な出力を定義する場合（例えば $y = \beta M$ ）

静特性：出力値をある理想的な値（ある値、ゼロ、大きな値など）に近づける場合
評価に用いる指標としては以下のようなものを考える。

SN 比 η 測定誤差の分散に対する有効な信号の変化の大きさに関係する値。大きい値を取るほど良い。10×対数スケールを用いる。

感度 S 有効な信号の変化の大きさに関係する量。問題により、大きい方、目標値なった方、小さい方がよいなど様々。10×対数スケールを用いる。

他の影響を排除したある制御因子の各水準の SN 比は？ → SN 比補助表

他の影響を排除したある制御因子の各水準の感度は？ → 感度補助表

10.2 プログラムの利用法

パラメータ設計のデータは図 1 のように、左の直交表の部分と右の実験結果の部分に分けられる。

| | E | F | G | H | 0.025 | 0.025 | 0.1 | 0.1 | 0.5 | 0.5 |
|---|---|---|---|---|-------|-------|-------|-------|-------|-------|
| 1 | 1 | 1 | 1 | 1 | 0.046 | 0.040 | 0.068 | 0.056 | 0.235 | 0.276 |
| 2 | 2 | 2 | 2 | 2 | 0.052 | 0.036 | 0.085 | 0.078 | 0.293 | 0.311 |
| 3 | 3 | 3 | 3 | 3 | 0.071 | 0.079 | 0.111 | 0.109 | 0.327 | 0.321 |
| 4 | 1 | 2 | 2 | 3 | 0.043 | 0.037 | 0.072 | 0.050 | 0.224 | 0.246 |
| 5 | 2 | 3 | 3 | 1 | 0.058 | 0.038 | 0.083 | 0.082 | 0.283 | 0.252 |
| 6 | 3 | 1 | 1 | 2 | 0.076 | 0.074 | 0.104 | 0.131 | 0.371 | 0.404 |
| 7 | 2 | 1 | 3 | 2 | 0.030 | 0.027 | 0.062 | 0.065 | 0.278 | 0.298 |
| 8 | 3 | 2 | 1 | 3 | 0.044 | 0.036 | 0.108 | 0.086 | 0.354 | 0.361 |
| 9 | 1 | 3 | 2 | 1 | 0.035 | 0.028 | 0.045 | 0.043 | 0.274 | 0.231 |

図 1 パラメータ設計のデータ

ここでは、制御因子が A～H の 8 種類、信号因子が 3 種類、誤差因子が 2 種類である。信号因子と誤差因子の部分の変数名には信号因子の数値が与えられている。

メニュー [分析－OR－品質管理－パラメータ設計] を選択すると図 2 のような分析メニューが表示される。

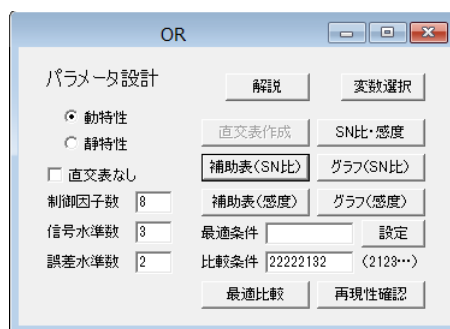


図 2 パラメータ設計分析メニュー

パラメータ設定には動特性と静特性の 2 種類あるが、今回は動特性のみにについて紹介する。まずメニュー中にある、「制御因子数」、「信号水準数」、「誤差水準数」の値を入力する。この例題の場合、デフォルトの数値がそのまま利用できる。次に「変数選択」ボタンですべての変数を選択する。現在の例では直交表が付いているが、単純に SN 比と感度のみを求める場合には、直交表を省略したデータを用いることもできる。その際には「直交表なし」チェックボックスにチェックを入れておく。

「SN 比・感度」ボタンをクリックすると、図 3 の計算結果が表示される。

| SN比・感度 | | | | | | | | | | | | |
|--------|---|---|---|---|---|---|---|---|---|--------|--------|--|
| | A | B | C | D | E | F | G | H | | SN比 | 感度 | |
| ▶ 1 | | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 26.534 | -5.730 | |
| 2 | | 1 | 1 | 2 | 2 | 2 | 2 | 2 | 2 | 28.258 | -4.239 | |
| 3 | | 1 | 1 | 3 | 3 | 3 | 3 | 3 | 3 | 23.149 | -3.516 | |
| 4 | | 1 | 2 | 1 | 1 | 2 | 2 | 3 | 3 | 26.656 | -6.431 | |
| 5 | | 1 | 2 | 2 | 2 | 3 | 3 | 1 | 1 | 25.083 | -5.230 | |

図 3 各実験の SN 比・感度

ここでは各実験に対して、単純に SN 比と感度を求めて表示している。

直交表を使った SN 比の補助表は「補助表 (SN 比)」ボタンをクリックすることで図 4 のように与えられる。感度の補助表については「補助表 (感度)」ボタンをクリックして得られる。

| 補助表 (SN比) | | | | |
|-----------|--------|--------|--------|---------|
| | 水準1 | 水準2 | 水準3 | MAX-MIN |
| ▶ A | 27.494 | 23.169 | | 4.325 |
| B | 23.656 | 22.961 | 29.378 | 6.418 |
| C | 28.332 | 25.959 | 21.704 | 6.628 |
| D | 22.358 | 26.353 | 27.284 | 4.926 |
| E | 26.649 | 24.186 | 25.160 | 2.463 |
| F | 25.040 | 27.077 | 23.878 | 3.199 |
| G | 26.675 | 24.008 | 25.313 | 2.667 |
| H | 23.489 | 26.686 | 25.820 | 3.197 |

図 4 補助表 (SN 比)

ここで制御因子 A は 2 水準であるから、空白が 1 つできている。

補助表をグラフにした図は「グラフ (SN 比)」ボタンをクリックして表示される。描画結果を図 5 に示す。

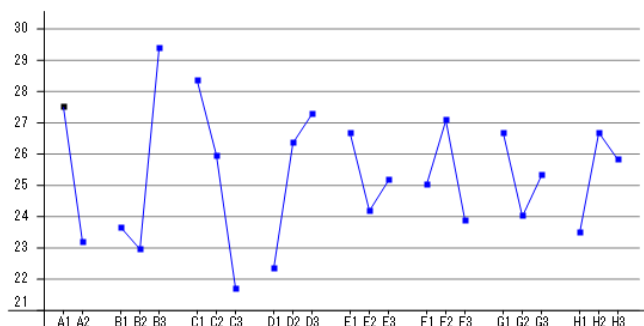


図 5 補助表のグラフ (SN 比)

図 4 と図 5 に対する感度の補助表とグラフは、それぞれ図 6 と図 7 で与えられる。

| | 水準 1 | 水準 2 | 水準 3 | MAX-MIN |
|---|--------|--------|--------|---------|
| A | -4.515 | -3.975 | | 0.540 |
| B | -3.458 | -5.041 | -4.236 | 1.583 |
| C | -4.337 | -3.729 | -4.669 | 0.940 |
| D | -6.000 | -4.117 | -2.618 | 3.381 |
| E | -3.963 | -4.762 | -4.011 | 0.799 |
| F | -3.309 | -4.239 | -5.187 | 1.878 |
| G | -4.045 | -4.400 | -4.290 | 0.355 |
| H | -5.052 | -3.968 | -3.716 | 1.336 |

図 6 補助表 (感度)

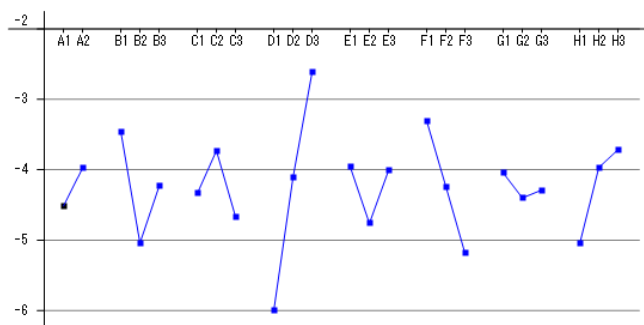


図 7 補助表のグラフ (感度)

SN 比の補助表やグラフを使った最適条件は「設定」ボタンをクリックすることでメニュー上の最適条件の部分に図 8 のように表示される。

| パラメータ設計 | | 解説 | 変数選択 |
|--------------------------------------|---|----------|-----------|
| <input checked="" type="radio"/> 動特性 | | 直交表作成 | SN比・感度 |
| <input type="radio"/> 静特性 | | 補助表(SN比) | グラフ(SN比) |
| <input type="checkbox"/> 直交表なし | | 補助表(感度) | グラフ(感度) |
| 制御因子数 | 8 | 最適条件 | 13131212 |
| 信号水準数 | 3 | 比較条件 | 22222132 |
| 誤差水準数 | 2 | | <2123...> |
| | | 最適比較 | 再現性確認 |

図 8 SN 比の最適条件の設定

比較条件で現在の実験から得られるデータの値を求めることができるが、最適条件との比較も可能である。これらの数値は「最適比較」ボタンで得ることができる。表示結果を図 9 に示す。最適条件の制御因子の組み合わせを変えることで結果を手動で訂正することもでき

る。

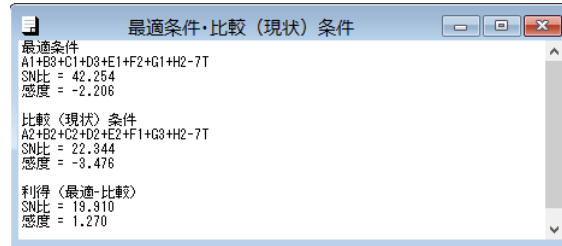


図 9 最適条件と比較条件

これらの最適条件と比較条件を実験で再現し、結果を得て、それをデータに追加する。当然その部分の直交表は空白になっているが、そのまま「再現性確認」ボタンをクリックすると図 10 に示す再現性確認表が得られる。

| 再現性確認表 | | | | |
|--------|--------|--------|--------|--------|
| | SN比推定値 | 感度推定値 | SN比実験値 | 感度実験値 |
| ▶ 最適条件 | 42.254 | -2.206 | 44.369 | -0.962 |
| 比較条件 | 22.344 | -3.476 | 26.177 | -4.059 |
| 利得 | 19.910 | 1.270 | 18.192 | 3.097 |

図 10 再現性確認表

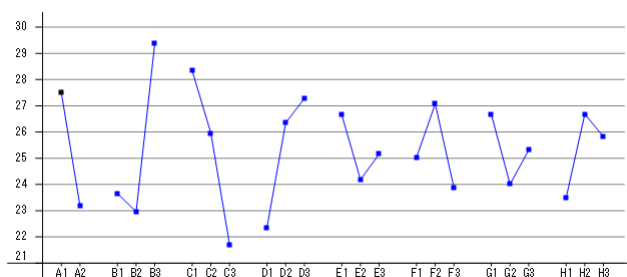
問題 1 動特性のパラメータ設計

パラメータ設計 1.txt の 2 頁目のデータを用いて以下の問いに答えよ。

- 1) 2 分類の制御因子が 1 つ、3 分類の制御因子が 7 つの場合の直交表がデータで与えたようになることを確認せよ。
- 2) 最初の 3 つの実験の SN 比と感度を求めよ。

| | SN 比 | 感度 |
|---|------|----|
| 1 | | |
| 2 | | |
| 3 | | |

- 3) 他の因子の影響を排除した制御因子 A の 2 つの水準の SN 比の値を求めよ。(補助表)
水準 1 [] 水準 2 []
- 4) 他の因子の影響を排除した制御因子 A の 2 つの水準の感度の値を求めよ。(補助表)
水準 1 [] 水準 2 []
- 5) SN 比の補助表による以下の要因効果図を表示せよ。最も変化の大きい制御因子は何か。



制御因子 []

6) 最も SN 比の良い制御因子の番号を書け。

A [] B [] C [] D [] E [] F [] G [] H []

7) SN 比の平均と最適値を求めよ。また最適値のときの感度を求めよ。

平均 [] 最適値 [] 感度 []

8) 現状で制御因子がすべて 2 の場合、SN 比と感度を求めよ。

SN 比 [] 感度 []

9) 最適値と現状との差を示せ。

SN 比 [] 感度 []

10) 制御因子をすべて最適値及び指定値にそろえて SN 比と感度の再現性を調べたところ、グリッドエディタの下から 2 行の数値になった。実験値に対する SN 比と感度を最適条件、比較条件、利得に分けて求めよ。

| | SN 比実験値 | 感度実験値 |
|------|---------|-------|
| 最適条件 | | |
| 比較条件 | | |
| 利得 | | |

11) 最適条件で与えた感度実験値の値-0.962 を 0 ($\beta=1$) に変更するように設定したい。

SN 値があまり大きく変化せず、感度が比較的大きく変化する制御因子として F に注目し、最適条件の 2 を 1 に変えた（これで感度の値は 0 に近づくはずである）。実験値から予想される SN 比と感度の変更された最適条件の値を求めよ。

| | SN 比実験値 | 感度実験値 |
|------|---------|-------|
| 最適条件 | | |

問題 2 静特性のパラメータ設計

パラメータ設計 2.txt のデータを用いて一般の望目特性（出力が安定していることが良い）として以下の問いに答えよ。（注：2,3,4）について Excel 解答の実験 4 が抜けている。）

1) 最初の 3 つの実験の SN 比と感度を求めよ。（信号水準 1，誤差水準 2）

| | SN 比 | 感度 |
|---|------|----|
| 1 | | |
| 2 | | |
| 3 | | |

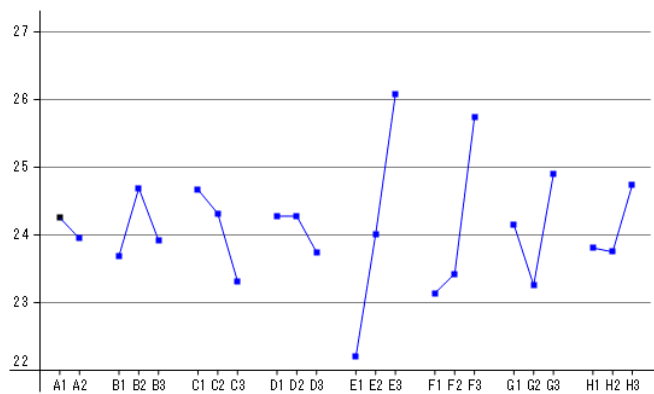
2) 他の因子の影響を排除した制御因子 B の 3 つの水準の SN 比の値を求めよ。（補助表）

水準 1 [] 水準 2 [] 水準 3 []

3) 他の因子の影響を排除した制御因子 B の 3 つの水準の感度の値を求めよ。（補助表）

水準 1 [] 水準 2 [] 水準 3 []

- 4) SN 比の補助表による以下の要因効果図を表示せよ。最も変化の大きい制御因子は何か。



制御因子 []

- 5) 設定条件があり、決まっているところもあるが、最適な制御因子の番号を埋めよ。
A [2] B [3] C [] D [2] E [] F [] G [] H [2]
- 6) SN 比の平均と最適値を求めよ。また最適値のときの感度を求めよ。
平均 [] 最適値 [] 感度 []
- 7) 現状で制御因子が A1B1C1D1E2F2G2H2 の場合、SN 比と感度を求めよ。
SN 比 [] 感度 []
- 8) 最適値と現状との差を示せ。
SN 比 [] 感度 []

問題 1 解答 動特性のパラメータ設計

- 2) 最初の 3 つの実験の SN 比と感度を求めよ。

| | SN 比 | 感度 |
|---|--------|--------|
| 1 | 26.534 | -5.730 |
| 2 | 28.258 | -4.239 |
| 3 | 23.149 | -3.516 |

- 3) 他の因子の影響を排除した制御因子 A の 2 つの水準の SN 比の値を求めよ。(補助表)
水準 1 [27.494] 水準 2 [23.169]
- 4) 他の因子の影響を排除した制御因子 A の 2 つの水準の感度の値を求めよ。(補助表)
水準 1 [-4.515] 水準 2 [-3.975]
- 5) SN 比の補助表による以下の要因効果図を表示せよ。最も変化の大きい制御因子は何か。
制御因子 [C]
- 6) 最も SN 比の良い制御因子の番号を書け。
A [1] B [3] C [1] D [3] E [1] F [2] G [1] H [2]
- 7) SN 比の平均と最適値を求めよ。また最適値のときの感度を求めよ。
平均 [25.332] 最適値 [42.254] 感度 [-2.206]
- 8) 現状で制御因子がすべて 2 の場合、SN 比と感度を求めよ。
SN 比 [23.076] 感度 [-4.516]
- 9) 最適値と現状との差を示せ。
SN 比 [19.178] 感度 [2.310]
- 10) 制御因子をすべて最適値及び指定値にそろえて SN 比と感度の再現性を調べたところ、グリッドエディタの下から 2 行の数値になった。実験値に対する SN 比と感度を最適条件、比較条件、利得に分けて求めよ。

| | SN 比実験値 | 感度実験値 |
|------|---------|--------|
| 最適条件 | 44.369 | -0.962 |
| 比較条件 | 26.177 | -4.059 |
| 利得 | 18.192 | 3.097 |

- 11) 最適条件で与えた感度実験値の値-0.962 を 0 ($\beta=1$) に変更するように設定したい。SN 値があまり大きく変化せず、感度が比較的大きく変化する制御因子として F に注目し、最適条件の 2 を 1 に変えた（これで感度の値は 0 に近づくはずである）。実験値から予想される SN 比と感度の変更された最適条件の値を求めよ。

| | SN 比実験値 | 感度実験値 |
|------|---------|-------|
| 最適条件 | | |

問題 2 解答 静特性のパラメータ設計

パラメータ設計 2.txt のデータを用いて一般の望目特性（出力が安定していることが良い）として以下の問いに答えよ。（注：2),3),4) について Excel 解答の実験 4 が抜けている。）

- 1) 最初の 3 つの実験の SN 比と感度を求めよ。（信号水準 1，誤差水準 2）

| | SN 比 | 感度 |
|---|--------|--------|
| 1 | 20.145 | 33.393 |
| 2 | 20.928 | 29.739 |
| 3 | 26.398 | 28.493 |

- 2) 他の因子の影響を排除した制御因子 B の 3 つの水準の SN 比の値を求めよ。（補助表）
水準 1 [23.695] 水準 2 [24.693] 水準 3 [23.921]
- 3) 他の因子の影響を排除した制御因子 B の 3 つの水準の感度の値を求めよ。（補助表）
水準 1 [29.825] 水準 2 [30.941] 水準 3 [29.907]
- 4) SN 比の補助表による以下の要因効果図を表示せよ。最も変化の大きい制御因子は何か。
制御因子 [E]
- 5) 設定条件があり、決まっているところもあるが、最適な制御因子の番号を埋めよ。
A [2] B [3] C [1] D [2] E [3] F [3] G [3] H [2]
- 6) SN 比の平均と最適値を求めよ。また最適値のときの感度を求めよ。
平均 [24.103] 最適値 [30.642] 感度 [30.458]
- 7) 現状で制御因子が A1B1C1D1E2F2G2H2 の場合、SN 比と感度を求めよ。
SN 比 [22.642] 感度 [29.537]
- 8) 最適値と現状との差を示せ。
SN 比 [8.000] 感度 [0.921]

10.3 パラメータ設計の理論

ここではまずゼロ点比例式の動特性パラメータ設計について考えるが、その前に SN 比と感度について、理論的な考察を加えておく。

1 つの実験では、信号水準 M_j ($j=1, \dots, p$) と誤差水準 N_α ($\alpha=1, \dots, n$) によって、表 1 のように pn 個のデータ $y_{j\alpha}$ が得られる。誤差水準はできるだけ広く散らばるよう配慮されるものとする。

表 1 パラメータ設計におけるデータ

| M_1 | | | ... | M_p | | |
|-------|-----|-------|-----|-------|-----|-------|
| N_1 | ... | N_n | ... | N_1 | ... | N_n |

| | | | | | | |
|----------|----------|----------|----------|----------|----------|----------|
| y_{11} | \cdots | y_{1n} | \cdots | y_{p1} | \cdots | y_{pn} |
|----------|----------|----------|----------|----------|----------|----------|

この実験についての、誤差水準 α のゼロ比例式回帰直線を考える。実測値 $y_{j\alpha}$ についての推定回帰式を $Y_{j\alpha} = b_\alpha M_j$ とすると、実測値との差の2乗和は以下となる。

$$EV_\alpha = \sum_{j=1}^p (y_{j\alpha} - Y_{j\alpha})^2 = \sum_{j=1}^p (y_{j\alpha} - b_\alpha M_j)^2$$

これを最小とするには、

$$\frac{\partial}{\partial b_\alpha} EV_\alpha = -2 \sum_{j=1}^p M_j (y_{j\alpha} - b_\alpha M_j) = 0$$

として、以下を得る。

$$b_\alpha = \sum_{j=1}^p M_j y_{j\alpha} / \sum_{j=1}^p M_j^2 = \frac{L_\alpha}{r}, \quad \text{ここに、} L_\alpha = \sum_{j=1}^p M_j y_{j\alpha}, \quad r = \sum_{j=1}^p M_j^2$$

全体のゼロ比例式回帰直線については、推定回帰式を $Y_j = b M_j$ とすると、実測値との差の2乗和は以下となる。

$$EV = \sum_{j=1}^p \sum_{\alpha=1}^n (y_{j\alpha} - Y_j)^2 = \sum_{j=1}^p \sum_{\alpha=1}^n (y_{j\alpha} - b M_j)^2$$

これを最小とするには、

$$\frac{\partial}{\partial b} EV = -2 \sum_{j=1}^p \sum_{\alpha=1}^n M_j (y_{j\alpha} - b M_j) = 0$$

として、以下を得る。

$$b = \sum_{j=1}^p \sum_{\alpha=1}^n M_j y_{j\alpha} / \sum_{j=1}^p \sum_{\alpha=1}^n M_j^2 = \frac{1}{nr} \sum_{j=1}^p \sum_{\alpha=1}^n M_j y_{j\alpha} = \frac{1}{nr} \sum_{\alpha=1}^n L_\alpha = \frac{1}{n} \sum_{\alpha=1}^n b_\alpha$$

次にデータの変動について考察する。まず、 $y=0$ からの全体の変動 S_T は以下となる。

$$S_T = \sum_{j=1}^p \sum_{\alpha=1}^n y_{j\alpha}^2 \quad \text{自由度 } pn$$

また、 $y=0$ からの全体の回帰変動 S_β は以下となる。

$$S_\beta = \sum_{j=1}^p \sum_{\alpha=1}^n (b M_j)^2 = nr b^2 = \frac{1}{nr} \left(\sum_{\alpha=1}^n L_\alpha \right)^2 \quad \text{自由度 } 1 \text{ (} b \text{ のみ)}$$

これより、 $b^2 = S_\beta / nr$ となる。

誤差水準 α の回帰直線の全体の回帰直線からの変動 $S_{N\beta}$ は以下となる。

$$S_{N\beta} = \sum_{j=1}^p \sum_{\alpha=1}^n (b_\alpha M_j - b M_j)^2 = r \sum_{\alpha=1}^n (b_\alpha - b)^2 = r \sum_{\alpha=1}^n b_\alpha^2 - nr b^2 = \sum_{\alpha=1}^n L_\alpha^2 / r - S_\beta$$

$$\text{自由度 } n-1 \quad \text{束縛条件} \quad \sum_{\alpha=1}^n (b_\alpha - b) = 0 \quad 1 \text{ 個}$$

各点の誤差水準 α の回帰直線からの変動 S_e は以下となる。

$$\begin{aligned}
 S_e &= \sum_{j=1}^p \sum_{\alpha=1}^n (y_{j\alpha} - b_\alpha M_j)^2 = \sum_{j=1}^p \sum_{\alpha=1}^n [y_{j\alpha} - (b_\alpha - b)M_j - bM_j]^2 \\
 &= \sum_{j=1}^p \sum_{\alpha=1}^n y_{j\alpha}^2 + \sum_{j=1}^p \sum_{\alpha=1}^n (b_\alpha - b)^2 M_j^2 + \sum_{j=1}^p \sum_{\alpha=1}^n b^2 M_j^2 - 2 \sum_{j=1}^p \sum_{\alpha=1}^n y_{j\alpha} b_\alpha M_j \\
 &= S_T + S_{NB} + S_\beta - 2 \sum_{\alpha=1}^n L_\alpha^2 / r^2 = S_T + S_{NB} + S_\beta - 2(S_{NB} + S_\beta) = S_T - S_{NB} - S_\beta \\
 &\quad \text{自由度 } pn - n \quad \text{束縛条件} \quad \sum_{j=1}^p M_j (y_{j\alpha} - b_\alpha M_j) = 0 \quad n \text{ 個}
 \end{aligned}$$

各点の誤差水準 α の回帰直線からの不偏分散 V_e は、変動を自由度で割って以下となる。

$$V_e = S_e / (np - n)$$

これが不偏分散となることはこの節の最後に詳しく説明する。

ここまでの議論で、全変動 S_T は、全体の回帰変動 S_β 、全体の回帰直線からの誤差水準 α の回帰変動 $S_{N\beta}$ 、各点の誤差水準 α の回帰直線からの変動 S_e の和で以下のように表されることが分かった。

$$S_T = S_\beta + S_{N\beta} + S_e$$

各点の全体の回帰直線からの変動 S_N は以下となる。

$$\begin{aligned}
 S_N &= \sum_{j=1}^p \sum_{\alpha=1}^n (y_{j\alpha} - bM_j)^2 = \sum_{j=1}^p \sum_{\alpha=1}^n y_{j\alpha}^2 + nb^2 \sum_{j=1}^p M_j^2 - 2b \sum_{j=1}^p \sum_{\alpha=1}^n y_{j\alpha} M_j \\
 &= S_T + nrb^2 - 2nrb^2 = S_T - S_\beta = S_{N\beta} + S_e \\
 &\quad \text{自由度 } pn - 1 \quad \text{束縛条件} \quad \sum_{j=1}^p \sum_{\alpha=1}^n M_j (y_{j\alpha} - bM_j) = 0
 \end{aligned}$$

各点の全体の回帰直線からの不偏分散 V_N は、変動を自由度で割って以下となる。

$$V_N = S_N / (pn - 1)$$

これらを使って、SN 比 η と感度 S を定義する。SN 比は、測定誤差の分散 σ^2 に対する有効な信号の変化の大きさ β^2 の比を用いて、また感度 S は β^2 の値を用いて以下のように定義される。

$$\text{SN 比 : } \eta = 10 \log_{10} \frac{\beta^2}{\sigma^2}, \quad \text{感度 : } S = 10 \log_{10} \beta^2$$

実際の計算では β^2 と σ^2 の値は不明であるので、これらの不偏推定量を用いて置き換える。

$$\text{SN 比 : } \eta = 10 \log_{10} \left[\frac{(S_\beta - V_e) / nr}{V_N} \right], \quad \text{感度 : } S = 10 \log_{10} [(S_\beta - V_e) / nr]$$

一般に SN 比は大きな値ほど、有効な信号を誤差の中から拾いやすくなり、良好な結果である。また、感度は対象により、大きな値がよい場合、小さな値がよい場合、目標値がよい場合など様々であるが、感度があまり変化しない制御因子を用いて SN 比を上げることを考え

る場合もある。

最後に、 β^2 の不偏推定量を求めておく。

$$y_{j\alpha} = \beta_\alpha M_j + \varepsilon_{j\alpha}, \quad E[\varepsilon_{j\alpha}] = 0, \quad E[\varepsilon_{j\alpha} \varepsilon_{j'\alpha'}] = \delta_{jj'} \delta_{\alpha\alpha'} V[\varepsilon]$$

とすると、

$$b_\alpha = \frac{1}{r} \sum_{j=1}^p M_j y_{j\alpha} = \frac{1}{r} \sum_{j=1}^p M_j (\beta_\alpha M_j + \varepsilon_{j\alpha}) = \beta_\alpha + \frac{1}{r} \sum_{j=1}^p M_j \varepsilon_{j\alpha}$$

より

$$\begin{aligned} S_e &= \sum_{j=1}^p \sum_{\alpha=1}^n (\beta_\alpha M_j + \varepsilon_{j\alpha} - b_\alpha M_j)^2 = \sum_{j=1}^p \sum_{\alpha=1}^n [(\beta_\alpha - b_\alpha) M_j + \varepsilon_{j\alpha}]^2 \\ &= \sum_{j=1}^p \sum_{\alpha=1}^n \left[-\frac{1}{r} M_j \sum_{j'=1}^p M_{j'} \varepsilon_{j'\alpha} + \varepsilon_{j\alpha} \right]^2 \\ &= \sum_{j=1}^p \sum_{\alpha=1}^n \left[\frac{1}{r^2} M_j^2 \left(\sum_{j'=1}^p M_{j'} \varepsilon_{j'\alpha} \right)^2 - \frac{2}{r} M_j \varepsilon_{j\alpha} \sum_{j'=1}^p M_{j'} \varepsilon_{j'\alpha} + \varepsilon_{j\alpha}^2 \right] \\ &= -\frac{1}{r} \sum_{\alpha=1}^n \sum_{j=1}^p \sum_{j'=1}^p M_j M_{j'} \varepsilon_{j\alpha} \varepsilon_{j'\alpha} + \sum_{j=1}^p \sum_{\alpha=1}^n \varepsilon_{j\alpha}^2 \end{aligned}$$

となり、

$$E[S_e] = E \left[-\frac{1}{r} \sum_{\alpha=1}^n \sum_{j=1}^p \sum_{j'=1}^p M_j M_{j'} \varepsilon_{j\alpha} \varepsilon_{j'\alpha} + \sum_{j=1}^p \sum_{\alpha=1}^n \varepsilon_{j\alpha}^2 \right] = (np - n) V(\varepsilon) \quad (\text{A1})$$

また、

$$b = \frac{\sum_{j=1}^p \sum_{\alpha=1}^n M_j y_{j\alpha}}{\sum_{j=1}^p \sum_{\alpha=1}^n M_j^2} = \frac{1}{nr} \sum_{j=1}^p \sum_{\alpha=1}^n M_j y_{j\alpha}$$

より、

$$\begin{aligned} b^2 &= \frac{1}{n^2 r^2} \left(\sum_{j=1}^p \sum_{\alpha=1}^n M_j (\beta_\alpha M_j + \varepsilon_{j\alpha}) \right)^2 \\ &= \frac{1}{n^2 r^2} \sum_{j=1}^p \sum_{\alpha=1}^n \sum_{j'=1}^p \sum_{\alpha'=1}^n M_j M_{j'} (\beta_\alpha \beta_{\alpha'} M_j M_{j'} + 2\beta_\alpha \varepsilon_{j\alpha} M_{j'} + \varepsilon_{j\alpha} \varepsilon_{j'\alpha'}) \\ &= \frac{1}{n^2 r^2} \left[\beta^2 n^2 r^2 + 2\beta n r \sum_{j=1}^p \sum_{\alpha=1}^n \varepsilon_{j\alpha} + \sum_{j=1}^p \sum_{\alpha=1}^n \sum_{j'=1}^p \sum_{\alpha'=1}^n M_j M_{j'} \varepsilon_{j\alpha} \varepsilon_{j'\alpha'} \right] \end{aligned}$$

であるから、

$$\begin{aligned}
 E[b^2] &= \frac{1}{n^2 r^2} E \left[\beta^2 n^2 r^2 + 2\beta n r \sum_{j=1}^p \sum_{\alpha=1}^n \varepsilon_{j\alpha} + \sum_{j=1}^p \sum_{\alpha=1}^n \sum_{j'=1}^p \sum_{\alpha'=1}^n M_j M_{j'} \varepsilon_{j\alpha} \varepsilon_{j'\alpha'} \right] \\
 &= \beta^2 + \frac{1}{n^2 r^2} \sum_{j=1}^p \sum_{\alpha=1}^n \sum_{j'=1}^p \sum_{\alpha'=1}^n M_j M_{j'} \delta_{jj'} \delta_{\alpha\alpha'} V(\varepsilon) \\
 &= \beta^2 + \frac{1}{nr} V(\varepsilon)
 \end{aligned}$$

よって、以下となる。

$$\beta^2 = E[b^2] - \frac{1}{nr} V(\varepsilon) \quad (\text{A2})$$

(A2)と(A1)、及び $b^2 = S_\beta / nr$ の関係から、

$$\beta^2 = \frac{1}{nr} E[S_\beta] - \frac{1}{nr} E[S_e / (np - n)] = E[(S_\beta - V_e) / nr] \quad (\text{A3})$$

すなわち、 β^2 の不偏推定量は $(S_\beta - V_e) / nr$ である。

同様の考え方で σ^2 の不偏推定量が $V_N = S_N / (pn - 1)$ であることも示すことができる。

次に我々は SN 比を最大にする制御因子の最適設定について考える。制御因子 A, B, …について直交表を作ると、他の制御因子の影響をならした、1つの制御因子の影響を調べることができるようになる。表 2 に直交表を加えたデータを示す。

表 2 パラメータ設計におけるデータ

| | A | B | … | M_1 | | | … | M_p | | | SN 比 | 感度 |
|---|---|---|---|-----------|---|-----------|---|-----------|---|-----------|----------|-------|
| | | | | N_1 | … | N_n | | N_1 | … | N_n | | |
| 1 | 1 | 1 | … | y_{111} | … | y_{11n} | … | y_{1p1} | … | y_{1pn} | η_1 | S_1 |
| 2 | 1 | 1 | … | y_{211} | … | y_{21n} | … | y_{2p1} | … | y_{2pn} | η_2 | S_2 |
| ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ |
| D | 2 | 2 | … | y_{d11} | … | y_{d1n} | … | y_{dp1} | … | y_{dpn} | η_d | S_d |

ここに SN 比と感度は上で述べた方法で求めて加えてあるものとする。直交表は、各制御因子の同じ番号の行を見ると、他の制御因子について、すべての番号が同じ数だけ入っているという特徴を持つ。

例えば制御因子 A が k になる行について、SN 比及び感度の平均を取ったものをそれぞれ $\bar{\eta}_{A=k}$ 、 $\bar{S}_{A=k}$ と書くとなると、SN 比の補助表は表 3 のようになる。感度の補助表も同様である。

表 3 SN 比の補助表

| 制御因子 | 水準 1 | … | 水準 r |
|------|--------------------|---|--------------------|
| A | $\bar{\eta}_{A=1}$ | … | $\bar{\eta}_{A=r}$ |

| | | | |
|---|--------------------|-----|--------------------|
| B | $\bar{\eta}_{B=1}$ | : | $\bar{\eta}_{B=r}$ |
| : | : | ... | : |

ここに水準の少ない制御因子の場合、その部分は空欄にしておく。

この補助表の SN 比の中で、制御因子ごとの水準値の最も大きな水準を並べたものを最適条件といい、例えば A1B2C1D3...などと表す。我々のプログラムでは制御因子名は省略して番号だけで表している。この最適な水準の SN 比を合計したものを SN 比の最適値という。感度についても SN 比の最適条件を用いて最適値を定義する。

これに対して現実の制御因子の設定を比較条件または現状条件という。この条件を用いて SN 比を合計したものを SN 比の比較値または現状値という。感度についても同様である。最適値と比較値の差は、今後の改善の可能性として検討すべき値である。

ここで述べた水準値は理論的な推測値である。この値が妥当なものかどうか、追実験をして検証しておかなければならない。また、現実的に考えて最適な制御条件が最良のものであるとは限らない。その際は、できるだけ SN 比の値を落とさず、感度で制御因子の調整を行うこともある。

参考文献

[1] 井上清和・中野恵司他, 入門パラメータ設計, 日科技連出版社, 2008.

11. オンライン品質工学

11.1 オンライン品質工学とは

工場の生産ラインでは、目標特性の製品を作るために、品質特性値を計測し、ラインを調整する。これらには時間と人手を要し、そこには費用が発生する。オンライン品質工学では、品質を金額ベースで表し、最適な計測間隔と調整間隔を求め、現行の値と比較し、どの程度の金額の削減につながるかを検討する。計算には以下の量を利用する。

目的特性の規格値： $m \pm \Delta$

不良品損失： A （円），計測コスト： B （円），調整コスト： C （円）

計測間隔： n_0 （個），調整限界： D_0 ，調整間隔： u_0 （個），計測タイムラグ： l （個）

バッチ処理の場合は、バッチ内標準偏差 s_m

全費用の最小化

$$L = \frac{B}{n} + \frac{C}{u} + \frac{A}{\Delta^2} \left[\frac{D^2}{3} + \left(\frac{n+1}{2} + l \right) \frac{D^2}{u} \right] \quad (\text{円/個})$$

仮定 $u = \lambda D^2$ ($\lambda = \text{const.}$)

11.2 プログラムの利用法

メニュー「分析－OR－品質管理－オンライン品質工学」を選択すると図 1 に示す分析画面が表示される。

品質管理

オンライン品質工学

実行選択

規格値 Δ 15 列数 4

不良品損失 A (円) 80 入力書式

計測コスト B (円) 150

調整コスト C (円) 1200

現行計測間隔 N_0 600

現行調整限界 D_0 5

現行調整間隔 U_0 1200

タイムラグ l 1

バッチ内標準偏差 s_m 0

時間当たり生産数 300

Clear 実行

注) 最適解は $U = \lambda D^2$ を仮定しています。
バッチ処理の場合、計測間隔等はバッチ単位です。

簡易計算

計算 計算結果

図 1 オンライン品質工学分析画面

このまま「実行」ボタンをクリックすると、画面上で与えられた数値を使って、図 2 に示す結果が表示される。



| | 現行 | 最適 |
|----------|----------|---------|
| ▶ 計測間隔N | 600.000 | 201.246 |
| 調整限界D | 5.000 | 3.811 |
| 調整間隔U | 1200.000 | 697.137 |
| パラメータ | 48.000 | 48.000 |
| 計測費/個 | 0.250 | 0.745 |
| 調整費/個 | 1.000 | 1.721 |
| 品質損失/個 | 5.196 | 2.478 |
| 総損失/個 | 6.446 | 4.944 |
| 工程能力指数Cp | 1.308 | 1.894 |
| RMS | 3.823 | 2.640 |
| バッチ内分散比率 | 0.000 | 0.000 |
| 累積計測回数 | 4.000 | 11.926 |
| 累積調整回数 | 2.000 | 3.443 |
| 累積計測費 | 600 | 1789 |
| 累積調整費 | 2400 | 4131 |
| 累積品質損失 | 12471 | 5947 |
| 累積総損失 | 15471 | 11867 |

図2 画面からの実行結果

累積となっている部分は、メニューの「時間当たり生産数」と「時間」の値を使って全時間で発生する回数と金額を計算している。例えば1日8時間稼働とすると「時間」を8にして、「時間当たり生産数」を1時間当たりの生産数にする。単位時間を1日にして、年間の稼働日数を「時間」として設定してもよい。

現行のデータを複数与えて、最適解を求めるときには、ファイルからチェックボックスをチェックして、変数選択して「実行」ボタンを押す。しかし、このプログラムは最適な結果を出すだけでなく、徐々に最適な結果に近づけて行くときにも利用できる。この方法を参考文献に従って問題形式で考えて行く。

オンライン品質工学 1.txt のファイルを開くと、図3のように現行のデータが入力されている。



| 製品1 | |
|-------------|------|
| 規格値Δ | 15 |
| 不良品損失A(円) | 80 |
| 計測コストB(円) | 150 |
| 調整コストC(円) | 1200 |
| 現行計測間隔N0(個) | 600 |
| 現行調整限界D0(個) | 5 |
| 現行調整間隔U0(個) | 1200 |
| タイムラグL(個) | 1 |
| バッチ内標準偏差Sm | 0 |
| 時間当たり生産数 | 300 |

1/3 (1.1) 分析: 備考:

図3 変更用ファイルデータ

データは1品目についてだけであるが、このデータを変更しながら最適な結果に近づけて行く。

質問1 最適計測間隔と最適調整限界を求めて現行より、いくらの改善かを推定する。

この問題の解答の最後には簡単な金額計算をする必要があるので、メニュー画面には下に簡単な電卓機能を付けておいた。

図3のデータで変数選択を「All」にして、「ファイルから」チェックボックスにチェックを入れ、「実行」ボタンをクリックすると、図4のような結果が出る。

| 計算結果 | | | | | | | | |
|----------|----------|---------|--------|--------|--------|--------|--------|--------|
| | 現行製品1 | 最適製品1 | 現行var2 | 最適var2 | 現行var3 | 最適var3 | 現行var4 | 最適var4 |
| ▶ 計測間隔N | 600.000 | 201.246 | | | | | | |
| 調整限界D | 5.000 | 3.811 | | | | | | |
| 調整間隔U | 1200.000 | 697.137 | | | | | | |
| パラメータs | 48.000 | 48.000 | | | | | | |
| 計測費/個 | 0.250 | 0.745 | | | | | | |
| 調整費/個 | 1.000 | 1.721 | | | | | | |
| 品質損失/個 | 5.196 | 2.478 | | | | | | |
| 総損失/個 | 6.446 | 4.944 | | | | | | |
| 工程能力指数Cp | 1.308 | 1.894 | | | | | | |
| RMS | 3.823 | 2.640 | | | | | | |
| バッチ内分散比率 | 0.000 | 0.000 | | | | | | |
| 累積計測回数 | 4.000 | 11.926 | | | | | | |
| 累積調整回数 | 2.000 | 3.443 | | | | | | |
| 累積計測費 | 600 | 1789 | | | | | | |
| 累積調整費 | 2400 | 4131 | | | | | | |
| 累積品質損失 | 12471 | 5947 | | | | | | |
| 累積総損失 | 15471 | 11867 | | | | | | |

図 4 ファイルからの実行結果

変数が現行と最適合わせて 4 つ表示されるのはすべての変数を選んでいないため、変数を 1 つずつ追加で選択して行く必要がないように、データの「規格Δ」の値が空欄のときは、関連する 2 列すべてを空欄になるように設定しているためである。もちろん他の列にデータがあるときには、その列も計算する。

最適な結果が出ているので、その値を参考にして、現行の値を変更することができる。グリッドエディタ上で、現行の値を次の列にコピーして必要な部分を変更した結果を図 5 に示す。

| データ編集 オンライン品質工学1.txt | | | | |
|----------------------|------|------|--|--|
| 製品1 | | | | |
| 規格値Δ | 15 | 15 | | |
| 不良品損失A(円) | 80 | 80 | | |
| 計測コストB(円) | 150 | 150 | | |
| 調整コストC(円) | 1200 | 1200 | | |
| 現行計測間隔N0(個) | 600 | 200 | | |
| 現行調整限界D0(個) | 5 | 4 | | |
| ▶ 現行調整間隔U0(個) | 1200 | 768 | | |
| タイムラグL(個) | 1 | 1 | | |
| バッチ内標準偏差Sm | 0 | 0 | | |
| 時間当たり生産数 | 300 | 300 | | |

図 5 現行値を変更した後のデータ

これを元に計算した結果が図 6 である。

| 計算結果 | | | | | | | | |
|----------|----------|---------|---------|---------|--------|--------|--------|--------|
| | 現行製品1 | 最適製品1 | 現行var2 | 最適var2 | 現行var3 | 最適var3 | 現行var4 | 最適var4 |
| ▶ 計測間隔N | 600.000 | 201.246 | 200.000 | 201.246 | | | | |
| 調整限界D | 5.000 | 3.811 | 4.000 | 3.811 | | | | |
| 調整間隔U | 1200.000 | 697.137 | 768.000 | 697.137 | | | | |
| パラメータs | 48.000 | 48.000 | 48.000 | 48.000 | | | | |
| 計測費/個 | 0.250 | 0.745 | 0.750 | 0.745 | | | | |
| 調整費/個 | 1.000 | 1.721 | 1.563 | 1.721 | | | | |
| 品質損失/個 | 5.196 | 2.478 | 2.648 | 2.478 | | | | |
| 総損失/個 | 6.446 | 4.944 | 4.961 | 4.944 | | | | |
| 工程能力指数Cp | 1.308 | 1.894 | 1.832 | 1.894 | | | | |
| RMS | 3.823 | 2.640 | 2.729 | 2.640 | | | | |
| バッチ内分散比率 | 0.000 | 0.000 | 0.000 | 0.000 | | | | |
| 累積計測回数 | 4.000 | 11.926 | 12.000 | 11.926 | | | | |
| 累積調整回数 | 2.000 | 3.443 | 3.125 | 3.443 | | | | |
| 累積計測費 | 600 | 1789 | 1800 | 1789 | | | | |
| 累積調整費 | 2400 | 4131 | 3750 | 4131 | | | | |
| 累積品質損失 | 12471 | 5947 | 6356 | 5947 | | | | |
| 累積総損失 | 15471 | 11867 | 11906 | 11867 | | | | |

図 6 現行値を変更した後の結果

ここでは現行の総損失の差を見てみる。1 個当たりの改善は 1.4850 円、1 時間 300 個、1 年 2000 時間として 89.1 万円となる。これはメニューの「時間」のところを「2000」に変えて出力すると、そのまま全時間の合計が計算されるので、差を見つけ易い。

工程能力指標 C_p を見てみると現行 1.308、改善 1.832 となる。

最後に、計測に必要なマンパワーを求めてみよう。計測に 3 分、調整に平均 15 分かるとする。計測回数は 1 日 8 時間として、1 日当たり 12 回、時間は $12 \times 3 = 36$ 分である。調整回数は 1 日当たり 3.443 回、時間は約 $3.443 \times 15 = 47.9$ 分である。両方合わせて、約 82.9 分となる。そのため、1 日当たり $82.9/480 = 0.173$ 人必要である。

規格区間が $[0, \Delta]$ となる場合については、図 1 の「 $[0, \Delta]$ 」ラジオボタンを選択し、その下の「中心」の位置を調整限界値の何倍に設定するかを決めて同じ処理を行うが、捜査は全く同じなので、説明は省略する。

問題 1

オンライン品質工学 1.txt のデータを用いて以下の問いに答えよ。これには単位時間当たりの生産数が含まれている。また、1 日当たりの就業時間を 8 時間とせよ。

- 1) 以下の値について現行値と最適値を求めよ。但し、最適値は少数でもよいとして、全費用をできるだけ少なくすること。

| | 現行 | 最適 | | 現行 | 最適 |
|-----------------|----|----|------------------|----|----|
| 計測間隔 N | | | 工程能力指数 | | |
| 調整限界 D | | | RMS (σ) | | |
| 調整間隔 U | | | 累積計測回数 | | |
| パラメータ λ | | | 累積調整回数 | | |
| 計測費/個 | | | 累積計測費 | | |
| 調整費/個 | | | 累積調整費 | | |
| 品質損失/個 | | | 累積品質損失 | | |
| 総損失/個 | | | 累積総損失 | | |

- 2) この結果より、計測間隔を 200、調整限界を 4 にする場合、調整間隔はいくらになるか。

$u = \lambda D^2$ の関係式を用いて計算せよ。 []

- 3) 新しい計測間隔、調整限界、調整間隔を入れて、新しいデータとして再計算して、新しい現行値を求めよ。

| | 新設定 | | 新設定 |
|--------|-----|------------------|-----|
| 計測間隔 N | | 工程能力指数 | |
| 調整限界 D | | RMS (σ) | |
| 調整間隔 U | | 累積計測回数 | |

| | | | |
|-----------------|--|--------|--|
| パラメータ λ | | 累積調整回数 | |
| 計測費/個 | | 累積計測費 | |
| 調整費/個 | | 累積調整費 | |
| 品質損失/個 | | 累積品質損失 | |
| 総損失/個 | | 累積総損失 | |

4) 現行と新しい設定とで、1 個当たりの総損失はどのように変化したか。

現行 [] 新設定 []

5) 現行と新しい設定とで、1 日当たりの総損失はどのように変化したか、いくら改善されたか。

現行 [] 新設定 [] 改善/日 []

6) 年間 2,000 時間 (1 日 8 時間) として現行と新しい設定とで、1 年当たりの総損失はどのように変化したか、いくら改善されたか。改善額 []

問題 2

オンライン品質工学 2.txt のデータを用いて以下の問いに答えよ。1 日当たりの就業時間を 8 時間とせよ。

以下の値について現行値と最適値を求めよ。但し、最適値は少数でもよいとして、全費用をできるだけ少なくすること。(時間当たりの製造数は 1 とする)

| | 現行 | 最適 | | 現行 | 最適 |
|-----------------|----|----|------------------|----|----|
| 計測間隔 N | | | 工程能力指数 | | |
| 調整限界 D | | | RMS (σ) | | |
| 調整間隔 U | | | 累積計測回数 | | |
| パラメータ λ | | | 累積調整回数 | | |
| 計測費/個 | | | 累積計測費 | | |
| 調整費/個 | | | 累積調整費 | | |
| 品質損失/個 | | | 累積品質損失 | | |
| 総損失/個 | | | 累積総損失 | | |

特に問題があるのは、不良品損失を与える A と品質損失を与える式 $u = \lambda D^2$ であろう。これらが自由に設定できる機能は必要かもしれない。

問題 1 解答

1) 以下の値について現行値と最適値を求めよ。但し、最適値は少数でもよいとして、全費用をできるだけ少なくすること。

| | 現行 | 最適 | | 現行 | 最適 |
|--------|------|---------|------------------|-------|--------|
| 計測間隔 N | 600 | 201.246 | 工程能力指数 | 1.308 | 1.894 |
| 調整限界 D | 5 | 3.811 | RMS (σ) | 3.823 | 2.640 |
| 調整間隔 U | 1200 | 691.137 | 累積計測回数 | 4.000 | 11.926 |

| | | | | | |
|-----------------|--------|--------|--------|-------|-------|
| パラメータ λ | 48.000 | 48.000 | 累積調整回数 | 2.000 | 3.443 |
| 計測費/個 | 0.250 | 0.745 | 累積計測費 | 600 | 1789 |
| 調整費/個 | 1.000 | 1.721 | 累積調整費 | 2400 | 4131 |
| 品質損失/個 | 5.196 | 2.478 | 累積品質損失 | 12471 | 5947 |
| 総損失/個 | 6.446 | 4.944 | 累積総損失 | 15471 | 11867 |

2) この結果より、計測間隔を 200、調整限界を 4 にする場合、調整間隔はいくらになるか。

$u = \lambda D^2$ の関係式を用いて計算せよ。 [768]

3) 新しい計測間隔、調整限界、調整間隔を入れて、新しいデータとして再計算して、新しい現行値を求めよ。

| | 新設定 | | 新設定 |
|-----------------|--------|------------------|--------|
| 計測間隔 N | 200 | 工程能力指数 | 1.832 |
| 調整限界 D | 4 | RMS (σ) | 2.729 |
| 調整間隔 U | 768 | 累積計測回数 | 12.000 |
| パラメータ λ | 48.000 | 累積調整回数 | 3.125 |
| 計測費/個 | 0.750 | 累積計測費 | 1800 |
| 調整費/個 | 1.563 | 累積調整費 | 3750 |
| 品質損失/個 | 2.648 | 累積品質損失 | 6356 |
| 総損失/個 | 4.961 | 累積総損失 | 11906 |

4) 現行と新しい設定とで、1 個当たりの総損失はどのように変化したか。

現行 [6.446] 新設定 [4.961]

5) 現行と新しい設定とで、1 日当たりの総損失はどのように変化したか、いくら改善されたか。

現行 [15471] 新設定 [11906] 改善/日 [3565]

6) 年間 2,000 時間 (1 日 8 時間) として現行と新しい設定とで、1 年当たりの総損失はどのように変化したか、いくら改善されたか。改善額 [891389]

問題 2 解答

以下の値について現行値と最適値を求めよ。但し、最適値は少数でもよいとして、全費用をできるだけ少なくすること。(時間当たりの製造数は 1 とする)

| | 現行 | 最適 | | 現行 | 最適 |
|-----------------|--------|-------|------------------|--------|--------|
| 計測間隔 N | 1.000 | 0.445 | 工程能力指数 | 0.554 | 0.882 |
| 調整限界 D | 8.000 | 3.749 | RMS (σ) | 5.416 | 3.402 |
| 調整間隔 U | 16.000 | 3.514 | 累積計測回数 | 8.000 | 17.984 |
| パラメータ λ | 0.250 | 0.250 | 累積調整回数 | 0.500 | 2.276 |
| 計測費/個 | 777 | 1747 | 累積計測費 | 6216 | 13.974 |
| 調整費/個 | 2021 | 9200 | 累積調整費 | 16166 | 73599 |
| 品質損失/個 | 57593 | 22727 | 累積品質損失 | 460740 | 181815 |
| 総損失/個 | 60390 | 33674 | 累積総損失 | 483122 | 269388 |

11.3 オンライン品質工学の理論

工場の生産ラインでは、目標特性の製品を作るために、品質特性値を計測し、ラインを調整する。これらには時間と人手を要し、そこには費用が発生する。オンライン品質工学では、品質を金額ベースで表し、最適な計測間隔と調整間隔を求め、現行の値と比較し、どの程度の金額の削減につながるかを検討する。

1) 目的特性の規格値が $[-\Delta, \Delta]$ の場合

計算には以下の量を利用するが、規格値の中央値は 0 に設定している。特にどんな値でもずらして考えればよい。

不良品損失： A ，計測コスト： B ，調整コスト： C

計測間隔： n ，調整間隔： u ，調整限界： D ，計測タイムラグ： l

製品 1 個当たりの計測コストは B/n 、調整コストは C/u である。また、製品 1 個当たりの品質損失（品質水準） Q を以下のように与える^[1]。

$$Q = \frac{A}{\Delta^2} \left[\frac{D^2}{3} + \left(\frac{n+1}{2} + l \right) \frac{D^2}{u} \right]$$

ここに、最初の項は調整限界内のばらつき、次の項は問題があった計測時の調整限界を超えた製品の予測個数と計測タイムラグによる計測が遅れた製品の個数である。この品質損失の角括弧の中は誤差分散であり、その平方根を誤差標準偏差または RMS と呼ぶ。

ここで与えた計測調整コストと品質損失を足して、製品 1 個当たりの総損失 L は以下となる。

$$L = \frac{B}{n} + \frac{C}{u} + \frac{A}{\Delta^2} \left[\frac{D^2}{3} + \left(\frac{n+1}{2} + l \right) \frac{D^2}{u} \right] \quad (1)$$

後の考察のために、品質損失（品質水準） Q を、損失関数という考え方に基づいて、少し補足的説明も加えて考えてみる。まず損失関数を以下と仮定する。

$$Q(x) = \frac{A}{\Delta^2} x^2$$

ここに係数は $Q(\Delta) = A$ となるように決めている。

今、損失を調整限界内損失（限界内損失と略す） Q_{in} と調整限界外損失（限界外損失と略す） Q_{out} に分けて考える。

限界内損失のデータの確率分布を、範囲 $-D \leq x \leq D$ の一様分布と考えると、確率密度関数は $1/2D$ となる。そのため損失関数の期待値は以下となる。

$$Q_{in} = \frac{1}{2D} \int_{-D}^D \frac{A}{\Delta^2} x^2 dx = \frac{A}{6D\Delta^2} [x^3]_{-D}^D = \frac{AD^2}{3\Delta^2}$$

次に限界外損失について考える。 n 個に 1 回の計測で管理外データが発見された場合、残りの $n-1$ 個については計測されていないのでエラーがあるかどうか分からない。そのため確率を $1/2$ として期待値をとると、 $(n-1)/2$ となる。これと管理外データが発見された 1 個を足して、期待値は $(n+1)/2$ となる。また、これにタイムラグ l による調整の遅れで管理外データが発生すると考えると、これらを合わせて、1 回の調整での管理外データの個数の期待値 n_{out} は以下となる。

$$n_{out} = \frac{n+1}{2} + l$$

これが調整間隔 u 個に 1 回起こるので、商品 1 個当たりの管理外データの発生確率 p_{out} は

以下となる。

$$p_{out} = \frac{n_{out}}{u} = \left(\frac{n+1}{2} + l \right) \frac{1}{u} \quad (2)$$

管理外データの損失関数の値は $x = D$ として、1 回につき AD^2/Δ^2 で与えられるので、商品 1 個当たりの限界外損失 Q_{out} は以下で与えられる。

$$Q_{out} = \frac{A}{\Delta^2} \left(\frac{n+1}{2} + l \right) \frac{D^2}{u}$$

限界内損失と限界外損失を合わせて品質損失は以下となる。

$$Q = Q_{in} + Q_{out} = \frac{A}{\Delta^2} \left[\frac{D^2}{3} + \left(\frac{n+1}{2} + l \right) \frac{D^2}{u} \right] \quad (3)$$

これは、分散という考えから求めた結果と同じである。これに計測費用と調整費用を加えて総費用は以下となる。

$$L = \frac{B}{n} + \frac{C}{u} + \frac{A}{\Delta^2} \left[\frac{D^2}{3} + \left(\frac{n+1}{2} + l \right) \frac{D^2}{u} \right]$$

現行計測間隔を n_0 、現行調整限界を D_0 、現行調整間隔を u_0 とすると、現行総損失 L_0 は以下のようになる。

$$L_0 = \frac{B}{n_0} + \frac{C}{u_0} + \frac{A}{\Delta^2} \left[\frac{D_0^2}{3} + \left(\frac{n_0+1}{2} + l \right) \frac{D_0^2}{u_0} \right] \quad (4)$$

次に、(4)式に基づく最適な総損失を求めてみよう。そのために、我々は $u = \lambda D^2$ という仮定を考える。パラメータ λ は、現行の値を使って $\lambda = u_0/D_0^2$ で与えられる。この関係は限界調整間隔を小さく設定すると、平均調整間隔も小さくする必要があるということに基づく。しかし、これには計算上の理由もある。(1)式にこの関係を代入すると以下となり、

$$L = \frac{B}{n} + \frac{C}{\lambda D^2} + \frac{A}{\Delta^2} \left[\frac{D^2}{3} + \left(\frac{n+1}{2} + l \right) \frac{1}{\lambda} \right] \quad (5)$$

n と D とは和の項として分離され、極小化条件を求めるために微分した後、単独に解を求めることが可能となる。他の関係では、解を求めるために、数値計算の必要が生じ、解析的には困難となる。

我々は(4)式を n と D とで微分して 0 とおき、最適計測間隔 \hat{n} と最適調整限界 \hat{D} を求める。これらは以下のように与えられる。

$$\hat{n} = \sqrt{\frac{2B\lambda\Delta^2}{A}}, \quad \hat{D} = \left(\frac{3C\Delta^2}{A\lambda} \right)^{1/4} \quad (6)$$

これより、最適総損失 \hat{L} は以下のようになる。

$$\hat{L} = \frac{B}{\hat{n}} + \frac{C}{\lambda \hat{D}^2} + \frac{A}{\Delta^2} \left[\frac{\hat{D}^2}{3} + \left(\frac{\hat{n}+1}{2} + l \right) \frac{1}{\lambda} \right] \quad (7)$$

ここで、ハットの付いた量を含む式は $u = \lambda D^2$ の仮定を入れていると考えてもらいたい。

次に $u = \lambda D^2$ の条件を外してその他の指標をみてみよう。誤差分散 σ^2 は(1)式の[]の中で、以下のように与えられる。

$$\sigma^2 = \frac{D^2}{3} + \left(\frac{n+1}{2} + l \right) \frac{D^2}{u}$$

誤差の標準偏差を RMS と呼び、以下で与えられる。

$$RMS = \sigma = \sqrt{\frac{D^2}{3} + \left(\frac{n+1}{2} + l \right) \frac{D^2}{u}} \quad (8)$$

また、規格値の範囲 $\pm\Delta$ とこの標準偏差の $\pm 3\sigma$ の範囲との比を工程能力指数 C_p と呼び、精度評価の1つの基準としている。

$$C_p = \frac{2\Delta}{6\sqrt{\frac{D^2}{3} + \left(\frac{n+1}{2} + l \right) \frac{D^2}{u}}} \quad (9)$$

バッチ処理の場合には、 A, n, u 等はバッチ単位の値として総損失等は計算される。計測方法としてバッチ内のいくつかの製品についてサンプリング検査されるので、バッチ処理の場合には計測のバッチ内分散 s_m^2 も考慮しなくてはならない。そのため、品質損失（品質水準） Q を以下のように考える。

$$Q = \frac{A}{\Delta^2} \left[\frac{D^2}{3} + \left(\frac{n+1}{2} + l \right) \frac{D^2}{u} + s_m^2 \right] \quad (10)$$

これに関連して、総損失 L や RMS、工程能力指数 C_p も変更を受ける。プログラムではバッチ内分散の値が全体の分散の値に占める比率をバッチ内分散比率 r_B として以下のように与えている。

$$r_B = \frac{s_B^2}{\frac{D^2}{3} + \left(\frac{n+1}{2} + l \right) \frac{D^2}{u} + s_m^2}$$

2) 目的特性の規格値が $[0, \Delta]$ の場合

規格の下限値は 0 に設定している。特にどんな値でもずらして考えればよい。ここでは、区間が正に限られた場合（製品によって負になる場合も含む）の損失関数を以下と考える。

$$Q(x) = \frac{A}{\Delta^2} x^2$$

限界内損失のデータの確率分布を、範囲 $m - D \leq x \leq m + D$ の一様分布と考えると、確率密度関数は $1/2D$ となる。そのため限界内損失関数の期待値は以下となる。

$$Q_{in} = \frac{1}{2D} \int_{m-D}^{m+D} \frac{A}{\Delta^2} x^2 dx = \frac{A}{6D\Delta^2} \left[x^3 \right]_{m-D}^{m+D} = \frac{A}{3\Delta^2} (D^2 + 3m^2)$$

定数 Δ の考え方が異なるので注意を要するが、形式的には以前の結果に比べて $3m^2$ だけ追加されている。

限界外損失の場合の損失関数の値は以下となる。)

$$\text{内側 (0 に近い側)} \quad Q(m-D) = A(m-D)^2 / \Delta^2$$

$$\text{外側 (0 から遠い側)} \quad Q(m+D) = A(m+D)^2 / \Delta^2$$

平均的に内側と外側で、確率 $1/2$ で管理外データが発生すると考えると、1節の p_{out} を用いて限界外損失 Q_{out} は以下ようになる。

$$\begin{aligned} Q_{out} &= \frac{1}{2} [Q(m-D) + Q(m+D)] p_{out} = \frac{A}{2\Delta^2} [(m+D)^2 + (m-D)^2] p_{out} \\ &= \frac{A}{\Delta^2} \left(\frac{n+1}{2} + l \right) \frac{D^2 + m^2}{u} \end{aligned}$$

ここで値が0未満になることも考えられるが、これは廃棄になることはないと考えて、損失関数がそのまま適用できると考える。廃棄になる場合は、それを見越して精度や中心の位置を調整する必要がある。

これより、品質損失は以下となる。

$$Q = Q_{in} + Q_{out} = \frac{A}{\Delta^2} \left[\frac{D^2 + 3m^2}{3} + \left(\frac{n+1}{2} + l \right) \frac{D^2 + m^2}{u} \right] \quad (11)$$

これに計測費用と調整費用を加えて総費用は以下となる。

$$L = \frac{B}{n} + \frac{C}{u} + \frac{A}{\Delta^2} \left[\frac{D^2 + 3m^2}{3} + \left(\frac{n+1}{2} + l \right) \frac{D^2 + m^2}{u} \right] \quad (12)$$

m が定数の場合（例えば Δ に比例するとしても）、1節で述べた、 n と D との和の項としての独立性は得られず、(4)式の最小化のためには連立方程式の数値計算が必要になる。しかし、 $m = \alpha D$ のように決めることができると、

$$L = \frac{B}{n} + \frac{C}{u} + \frac{A}{\Delta^2} \left[\frac{(1+3\alpha^2)D^2}{3} + \left(\frac{n+1}{2} + l \right) \frac{(1+\alpha^2)D^2}{u} \right]$$

のようになり、前と同じ方法で計算が容易になる。

今、 $u = \lambda D^2$ を仮定すると

$$L = \frac{B}{n} + \frac{C}{\lambda D^2} + \frac{A}{\Delta^2} \left[\frac{(1+3\alpha^2)D^2}{3} + \left(\frac{n+1}{2} + l \right) \frac{1+\alpha^2}{\lambda} \right] \quad (13)$$

我々は(4)式を n と D とで微分して0とおき、最適計測間隔 \hat{n} と最適調整限界 \hat{D} を求めると、以下のように与えられる。

$$\hat{n} = \sqrt{\frac{2B\lambda\Delta^2}{A(1+\alpha^2)}} \quad , \quad \hat{D} = \left(\frac{3C\Delta^2}{A\lambda(1+3\alpha^2)} \right)^{1/4} \quad (14)$$

誤差分散 σ^2 については、損失関数に関係なく、(1)式の[]の中で、以下のように与えられる。

$$\sigma^2 = \frac{D^2}{3} + \left(\frac{n+1}{2} + l \right) \frac{D^2}{u}$$

誤差の標準偏差を RMS も同様である。

$$RMS = \sigma = \sqrt{\frac{D^2}{3} + \left(\frac{n+1}{2} + l \right) \frac{D^2}{u}} \quad (15)$$

工程能力指数 C_p については目的特性の範囲が異なるため、以下とする。

$$C_p = \frac{\Delta}{6 \sqrt{\frac{D^2}{3} + \left(\frac{n+1}{2} + l \right) \frac{D^2}{u}}} \quad (16)$$

バッチ処理の場合には、前と同様に計測のバッチ内分散 s_m^2 も考慮して、品質損失（品質水準） Q を以下のように考える。

$$Q = \frac{A}{\Delta^2} \left[\frac{D^2 + 3m^2}{3} + \left(\frac{n+1}{2} + l \right) \frac{D^2 + m^2}{u} + s_m^2 \right] \quad (17)$$

$u = \lambda D^2$ と $m = \alpha D$ の仮定の下に総費用は以下となる。

$$L = \frac{B}{n} + \frac{C}{\lambda D^2} + \frac{A}{\Delta^2} \left[\frac{(1+3\alpha^2)D^2}{3} + \left(\frac{n+1}{2} + l \right) \frac{1+\alpha^2}{\lambda} + s_m^2 \right] \quad (18)$$

バッチ内分散の値は定数として与えるので、(14)式で与えられる最適計測間隔 \hat{n} と最適調整限界 \hat{D} には変更がない。

参考文献

- [1] 矢野宏, 品質工学概論, 日本規格協会, 2009.

12. 異常検知

12.1 異常検知とは

製造現場における検査過程では多くのデータが測定されるが、正常なデータと異常なデータの迅速な選別は品質管理の上で非常に重要である。ここではその主要な理論について説明し、それを実践するプログラムを紹介する。理論はデータが多変量正規分布に従うと仮定される場合とそうでない場合を扱う。データが多変量正規分布に従う場合、マハラノビス距離の2乗を元にしたホテリングの t^2 統計量に基づく判定法を用いる。また、多変量正規分布に従わない場合は、確率的な解釈も可能な混合正規分布モデルを仮定する方法を用いている。また、1次元データについては、ガンマ分布による異常検知の方法も加えている。

12.2 プログラムの利用法

メニュー「分析－OR－品質管理－異常検知」を選択すると、図1のような分析実行画面が表示される。

図1 異常検知の分析実行画面

このプログラムでは、グリッドエディタに表示されているデータから、異常データを選別する「データクレンジング」機能と、別頁にある正常データを用いて、現在表示されているデータの異常データを選別する「正常データとの比較」機能がある。後者の場合は、正常データがどの頁にあるかを指定しなければならない。

分析には、(多変量)「正規分布からの検知」、(多変量)「混合正規分布からの検知」、「時系列データの検知」、「入出力データの検知」がある。最初の「正規分布からの検知」は1つの正規分布からのデータのずれを検知するもので、マハラノビス距離を基にした手法である。これは分布が限定されているが、多くのデータでほぼ正規分布の仮定が成り立つと考えられるので、利用範囲は広い。「混合正規分布からの検知」は、正規分布が仮定できない場合で、しかもどのような分布になっているか予想が困難な場合に適用が可能である。これは、分布を複数の正規分布の重ね合わせとして考えるモデルで、いくつかの正規分布の重ね合わせで考えると効果的かという判断も可能である。「時系列データの検知」では、最も実用的と思われる「特異SP変換法」(特異スペクトル変換法)が利用できる。「入出力データの検知」で

は、「重回帰」分析を基本として、多重共線性のある場合の手法で「リッジ回帰」分析と「PLS 回帰」分析が利用可能である。

他に「1次元ガンマ分布からの検知」があるが、これは変数が1つの場合にしか適用できないので、現実には利用しにくいかも知れない。以後、まず2変数のデータを用いて、正規分布と混合分布からの検知について、プログラムの説明をする。

図2に示すファイル「異常検知1(正規分布).txt」の3ページ目は、2変数で異常値を含んだデータである。

| | 身長 | 体重 |
|----|-----|----|
| 1 | 148 | 41 |
| 2 | 160 | 49 |
| 3 | 159 | 45 |
| 4 | 153 | 43 |
| 5 | 151 | 42 |
| 6 | 140 | 29 |
| 7 | 158 | 49 |
| 8 | 137 | 31 |
| 9 | 149 | 47 |
| 10 | 160 | 47 |
| 11 | 151 | 42 |

図2 異常値を含んだデータ

この中から異常値を検出するには、「データクレンジング」ラジオボタンを選び、「確率/分位点」ラジオボタンを選んで、異常値の確率値を指定する（この場合は確率値となる）。ここでは5%に設定している。その後、正規分布からの異常検知の「検知」ボタンをクリックすると図3のような出力結果を得る。

| | Maha2乗 | 異常度(f値) | 確率 | 異常 |
|----|--------|---------|-------|----|
| 15 | 2.011 | 0.908 | 0.415 | 0 |
| 16 | 2.851 | 1.288 | 0.292 | 0 |
| 17 | 0.658 | 0.297 | 0.745 | 0 |
| 18 | 1.026 | 0.463 | 0.634 | 0 |
| 19 | 0.565 | 0.255 | 0.776 | 0 |
| 20 | 1.789 | 0.808 | 0.456 | 0 |
| 21 | 14.438 | 6.520 | 0.005 | 1 |
| 22 | 1.718 | 0.776 | 0.470 | 0 |
| 23 | 3.303 | 1.492 | 0.242 | 0 |
| 24 | 0.310 | 0.140 | 0.870 | 0 |

図3 データクレンジング検知結果

出力は、このデータから求めた多変量正規分布の平均からのマハラノビス距離の2乗、それを元にしたホテリング統計量（F分布のf値）、その検定確率、異常かどうかの判定である。判定は正常と異常でそれぞれ0または1で出力される。

次に、「統計量」ボタンをクリックすると、図4のように、平均や共分散等のパラメータ推定値等と共に、異常の判定に使われるホテリング統計量の閾値が出力される。利用者はこの値を参考にして、閾値としてホテリング統計量を用いてもよい。

| | 身長 | 体重 |
|------------|---------|--------|
| ▶ 平均 | 149.000 | 38.700 |
| 分散共分散 | | |
| 身長 | 51.733 | 39.433 |
| 体重 | 39.433 | 40.343 |
| 閾値確率 | 0.050 | |
| 自由度 | 2.28 | |
| 閾値(Hit-f値) | 3.340 | |

図4 パラメータ推定値

次に「正常データとの比較」ラジオボタンをクリックし、同じファイルの1頁目を開く、正常データは2頁目に入っているものとして、「正常データ」テキストボックスの中に2を入力する。「検知」ボタンをクリックすると、図5のように2頁目の正常データから求められるパラメータを元にした、異常検知結果が出力される。

| | Maha2乗 | 異常度(f値) | 確率 | 異常 |
|-----|--------|---------|-------|----|
| ▶ 1 | 2.295 | 1.036 | 0.368 | 0 |
| 2 | 0.484 | 0.219 | 0.805 | 0 |
| 3 | 10.957 | 4.948 | 0.014 | 1 |
| 4 | 0.368 | 0.166 | 0.848 | 0 |
| 5 | 2.384 | 1.077 | 0.354 | 0 |
| 6 | 0.645 | 0.292 | 0.749 | 0 |
| 7 | 2.763 | 1.248 | 0.303 | 0 |
| 8 | 2.048 | 0.925 | 0.408 | 0 |
| 9 | 2.412 | 1.089 | 0.350 | 0 |
| 10 | 6.417 | 2.898 | 0.072 | 0 |

図5 正常データとの比較検知結果

出力項目については図3と同様である。「統計量」ボタンをクリックした結果は、正常データを元にした結果であり、図4と同じ様式であるので省略する。

ファイル「異常検知 3(複合正規分布).txt」を図6のように読み込み、「データクレンジング」ラジオボタンを選択し、データ処理の結果を見てみる。

| | 変数1 | 変数2 |
|-----|-------|-------|
| ▶ 1 | 10.19 | 4.32 |
| 2 | 9.58 | 1.48 |
| 3 | 9.71 | 4.85 |
| 4 | 11.49 | 5.35 |
| 5 | 10.59 | 10.92 |
| 6 | 9.15 | 0.09 |
| 7 | 8.75 | 4.96 |
| 8 | 9.87 | 7.96 |
| 9 | 10.04 | 7.99 |
| 10 | 8.99 | 6.81 |

図6 非正規分布のデータ

このデータは、実際には2つの正規分布を合わせたものであるが、今の段階ではそれが分からないものとする。仮に「混合」テキストボックスを2とし、処理を進める。後にこの数字を変更して最も良いモデルを選択する。

「分類」ボタンをクリックすると、図7のように、2つの群についてのデータの帰属度と分類結果が得られる。

| | 帰属度1 | 帰属度2 | 分類 |
|-----|-------|-------|----|
| 343 | 0.999 | 0.001 | 1 |
| 344 | 1.000 | 0.000 | 1 |
| 345 | 0.999 | 0.001 | 1 |
| 346 | 0.984 | 0.016 | 1 |
| 347 | 0.999 | 0.001 | 1 |
| 348 | 0.999 | 0.001 | 1 |
| 349 | 0.999 | 0.001 | 1 |
| 350 | 1.000 | 0.000 | 1 |
| 351 | 0.000 | 1.000 | 2 |
| 352 | 0.000 | 1.000 | 2 |

図7 レコード毎の帰属度と分類結果

この分類結果をコピーして、元のデータに図8のように貼り付け、

| データ編集 異常検知3(複合正規分布).txt | | | |
|-------------------------|-------|-------|---|
| | 変数1 | 変数2 | |
| 1 | 10.19 | 4.32 | 1 |
| 2 | 9.58 | 1.48 | 1 |
| 3 | 9.71 | 4.85 | 1 |
| 4 | 11.49 | 5.35 | 1 |
| 5 | 10.59 | 10.92 | 1 |
| 6 | 9.15 | 0.09 | 1 |
| 7 | 8.75 | 4.96 | 1 |
| 8 | 9.87 | 7.96 | 1 |
| 9 | 10.04 | 7.99 | 1 |
| 10 | 8.99 | 6.81 | 1 |

図8 分類項目の貼り付け

分析「相関と回帰分析」の「先頭列で群分け」ラジオボタンを選択して、図9のような散布図を描くことも可能である。

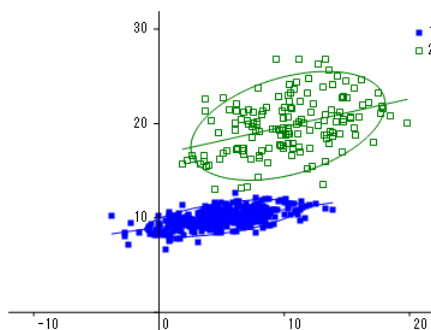


図9 データの散布図

但し、このグラフには、グラフメニュー「設定－正規楕円半径」を用いて、 2σ の確率楕円を加えてある。

次に「統計量」ボタンをクリックすると、図10のように2つの群の平均と共分散、群の生起確率の推測値などが表示される。この中で一番下のBICはモデル判定によく利用されるベイズ情報量基準と呼ばれるもので、この値が小さいほど良いモデルとして評価される。

| | 変数1 | 変数2 |
|--------|-----------|--------|
| 群1 | | |
| 生起確率 | 0.697 | |
| 平均 | 9.973 | 5.091 |
| 共分散 | | |
| 変数1 | 0.992 | 1.845 |
| 変数2 | 1.845 | 9.744 |
| 群2 | | |
| 生起確率 | 0.303 | |
| 平均 | 19.738 | 10.267 |
| 共分散 | | |
| 変数1 | 8.540 | 4.666 |
| 変数2 | 4.666 | 15.133 |
| 閾値(a値) | 7.305 | |
| 対数尤度 | -2392.768 | |
| AIC | 4809.535 | |
| BIC | 4860.110 | |

図 10 2 群の場合のパラメータ推定値

現在の 2 群の場合は BIC=4860.110 であるが、3 群にすると 4888.488 となり、2 群の方が良いモデルであると判断される。これは、2 群を故意に作ったモデルであるので、当然の結果である。

また、「検知」ボタンをクリックすると、図 11 のように異常度の値と判別結果が表示される。

| | 異常度(a値) | 異常 |
|-----|---------|----|
| 97 | 5.259 | 0 |
| 98 | 4.859 | 0 |
| 99 | 4.379 | 0 |
| 100 | 3.285 | 0 |
| 101 | 3.165 | 0 |
| 102 | 3.323 | 0 |
| 103 | 7.567 | 1 |
| 104 | 3.786 | 0 |
| 105 | 3.606 | 0 |
| 106 | 3.473 | 0 |

図 11 混合正規分布からの検知

判定には「確率/分位点」ラジオボックスの分位点を利用している。

1 次元ガンマ分布でも同様の結果の表示となるので、ここでは省略するが、ガンマ分布の場合は、2 つのパラメータが推測される。

時系列データからの異常検知では、正常データを指定しないので、図 1 の分析メニューでは「データクレンジング」ラジオボタンを選んで実行する。時系列データは、図 12 のように与えられる。

| | data |
|----|------|
| 1 | 57.8 |
| 2 | 61.8 |
| 3 | 38.8 |
| 4 | 8.8 |
| 5 | 1.4 |
| 6 | 24.7 |
| 7 | 50.5 |
| 8 | 63.0 |
| 9 | 47.2 |
| 10 | 13.8 |

図 12 時系列データ (異常検知 4(特異 SP 変換).txt)

異常検知の方法は簡単で、変数選択をして、「特異 SP 変換」ボタンをクリックする。その際、必要ならば「時刻数」、「系列数」、「ラグ」、「成分数」の値を変更する。これらの値はそれぞれ理論の説明の中の、 w, k, L, m に相当する。「目盛」の値は、グラフの目盛間隔の値を与える。図 13 に実行結果を示す。

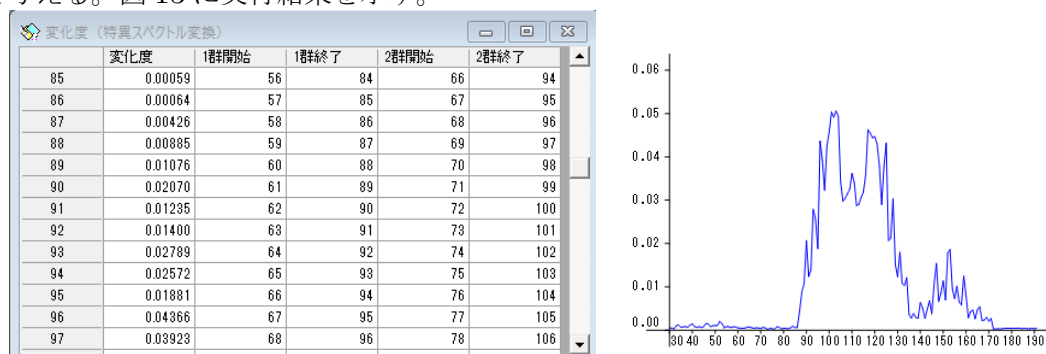


図 13 特異スペクトル変換による異常検知結果

図 13 の表の中で、左端が時刻であり、右 4 列は 2 つの群が利用したデータの範囲を示している。1 群は時刻の 1 つ前までのデータを利用し、2 群はそれから「ラグ」だけ遅れた時刻のデータを使っている。

入力と出力があるデータでは、重回帰分析が基本となる。通常の重回帰分析と同様に、目的変数、説明変数の順に変数を選んで、図 1 の分析メニューの「重回帰分析」ボタンをクリックすると図 14 のような実行結果が表示される。

| 異常度 (重回帰分析) | | | | | |
|-------------|-----|--------|--------|-------|----|
| | 実測値 | 予測値 | 残差 | 異常度 | 判定 |
| 4 | 39 | 40.859 | -1.859 | 0.160 | 0 |
| 5 | 81 | 84.237 | -3.237 | 0.486 | 0 |
| 6 | 47 | 53.870 | -6.870 | 2.192 | 0 |
| 7 | 92 | 91.659 | 0.341 | 0.005 | 0 |
| 8 | 75 | 76.076 | -1.076 | 0.054 | 0 |
| 9 | 77 | 79.557 | -2.557 | 0.304 | 0 |
| 10 | 64 | 61.139 | 2.861 | 0.380 | 0 |
| 11 | 76 | 82.910 | -6.910 | 2.217 | 0 |
| 12 | 97 | 85.622 | 11.378 | 6.011 | 1 |
| 13 | 65 | 65.551 | -0.551 | 0.014 | 0 |

図 14 重回帰分析実行結果 (異常検知 5(入出力).txt 1 頁目)

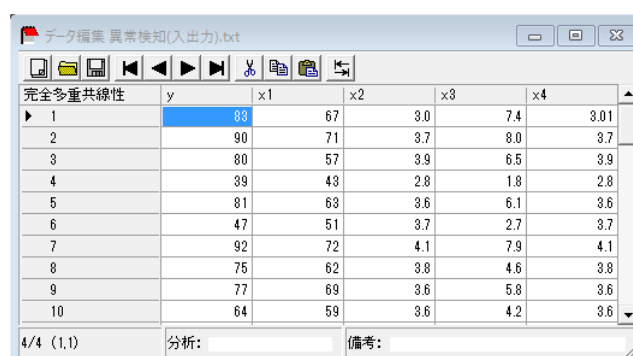
また、隣の「統計量」ボタンをクリックすると、図 15 のような結果が表示される。

| 統計量 (重回帰分析) | | | | | | |
|-------------|--------|--------|--------|---------|--------|-------|
| | 偏回帰係数 | 標準化係数 | VIF | 残差分散 | 閾値 | 重相関R |
| ▶ x1 | 0.1490 | 0.0749 | 1.1796 | 21.5362 | 4.0705 | 0.943 |
| x2 | 2.2283 | 0.0527 | 1.0554 | | | |
| x3 | 2.7614 | 0.3522 | 3.9544 | | | |
| x4 | 0.4314 | 0.5888 | 3.9670 | | | |
| 切片 | 6.8654 | 0.0000 | | | | |

図 15 重回帰分析統計量 (異常検知 5(入出力).txt 1 頁目)

これは正常に重回帰分析が行われた結果である。

データに多重共線性がある場合、例えば、図 16 は変数 x_2 と x_4 が殆ど同じで、1 番目のデータだけが、0.01 違っている例である。



| 完全多重共線性 | y | x1 | x2 | x3 | x4 |
|---------|----|----|-----|-----|------|
| ▶ 1 | 83 | 67 | 3.0 | 7.4 | 3.01 |
| 2 | 90 | 71 | 3.7 | 8.0 | 3.7 |
| 3 | 80 | 57 | 3.9 | 6.5 | 3.9 |
| 4 | 39 | 43 | 2.8 | 1.8 | 2.8 |
| 5 | 81 | 63 | 3.6 | 6.1 | 3.6 |
| 6 | 47 | 51 | 3.7 | 2.7 | 3.7 |
| 7 | 92 | 72 | 4.1 | 7.9 | 4.1 |
| 8 | 75 | 62 | 3.8 | 4.6 | 3.8 |
| 9 | 77 | 69 | 3.6 | 5.8 | 3.6 |
| 10 | 64 | 59 | 3.6 | 4.2 | 3.6 |

図 16 強い多重共線性があるデータ（同 4 頁目）

このデータに対して重回帰分析の「統計量」ボタンをクリックすると、図 17 のような結果が得られる。



| | 偏回帰係数 | 標準化係数 | 残差分散 | 閾値 | 重相関R | 寄与率R ² |
|------|-----------|----------|---------|--------|-------|-------------------|
| ▶ x1 | 0.2251 | 0.1132 | 38.0310 | 3.3607 | 0.897 | 0.804 |
| x2 | 521.1502 | 12.3456 | 共線性指標 | 0.0000 | | |
| x3 | 6.7407 | 0.8596 | | | | |
| x4 | -519.5957 | -12.2978 | | | | |
| 切片 | 18.5670 | 0.0000 | | | | |

図 17 強い多重共線性があるデータの重回帰分析統計量（同 4 頁目）

この結果ではほぼ同じ値である変数 x2 と x4 の係数が極端に大きくなっている。これは予測において、x2 と x4 の値が少しずれると大きな差となって表れることを示しており、予測の頑健性において問題となる。

これを解決するための代表的な手法に、リッジ回帰分析と PLS 回帰分析がある。理論のところで示したように、リッジ回帰分析は多重共線性の元となる説明変数の共分散行列の対角成分に、ある定数（ここでは、 $\eta \div N$ ）を加える方法である。この定数 η は、1 個抜き交差検証の残差分散が最小となるように決める。

図 16 のデータに対して、「リッジ回帰分析」とその「統計量」を与えた結果を図 18 と図 19 に示す。



| | 実測値 | 予測値 | 残差 | 異常度 | 判定 |
|----|-----|--------|--------|-------|----|
| 4 | 39 | 41.477 | -2.477 | 0.267 | 0 |
| 5 | 81 | 84.407 | -3.407 | 0.505 | 0 |
| 6 | 47 | 53.606 | -6.606 | 1.899 | 0 |
| 7 | 92 | 90.447 | 1.553 | 0.105 | 0 |
| 8 | 75 | 76.297 | -1.297 | 0.073 | 0 |
| 9 | 77 | 79.579 | -2.579 | 0.290 | 0 |
| 10 | 64 | 60.639 | 3.361 | 0.492 | 0 |
| 11 | 76 | 82.966 | -6.966 | 2.112 | 0 |
| 12 | 97 | 84.711 | 12.289 | 6.573 | 1 |
| 13 | 65 | 65.043 | -0.043 | 0.000 | 0 |

図 18 リッジ回帰分析による異常検知



| | 偏回帰係数 | 標準化係数 | 残差分散 | 閾値 | 重相関R | 寄与率R ² |
|------|--------|--------|---------|--------|-----------|-------------------|
| ▶ x1 | 0.1693 | 0.0851 | 22.9761 | 3.9367 | 0.942 | 0.887 |
| x2 | 0.8338 | 0.0198 | | | 最良 η | 8.600 |
| x3 | 2.2780 | 0.2905 | | | | |
| x4 | 0.4691 | 0.6402 | | | | |
| 切片 | 9.9390 | 0.0000 | | | | |

図 19 リッジ回帰分析統計量

また、PLS 回帰分析は多重共線性に対して、独立成分として与える変数の数を減らす方

法で対応する。どれだけの変数を減らすかは、リッジ回帰分析と同様に 1 個抜き交差検証の残差分散が最小となるように決める。図 20 と図 21 にメニューの指定で独立成分を 3 にした PLS 回帰分析を実行した結果を示す。



| | 実測値 | 予測値 | 残差 | 異常度 | 判定 |
|----|-----|--------|--------|-------|----|
| 4 | 39 | 41.067 | -2.067 | 0.196 | 0 |
| 5 | 81 | 83.798 | -2.798 | 0.359 | 0 |
| 6 | 47 | 54.035 | -7.035 | 2.268 | 0 |
| 7 | 92 | 92.300 | -0.300 | 0.004 | 0 |
| 8 | 75 | 75.417 | -0.417 | 0.008 | 0 |
| 9 | 77 | 79.388 | -2.388 | 0.261 | 0 |
| 10 | 64 | 61.694 | 2.306 | 0.244 | 0 |
| 11 | 76 | 82.522 | -6.522 | 1.949 | 0 |
| 12 | 97 | 85.634 | 11.366 | 5.921 | 1 |
| 13 | 65 | 66.501 | -1.501 | 0.103 | 0 |

図 20 PLS 回帰分析による異常検知



| | 偏回帰係数 | 標準化係数 | r-VIF | 残差分散 | 閾値 | 重相関R | 寄与率R^2 |
|----|--------|--------|--------|---------|--------|-------|--------|
| x1 | 0.1312 | 0.0659 | 1.0160 | 21.8210 | 4.9222 | 0.942 | 0.888 |
| x2 | 2.7116 | 0.0642 | 1.1392 | 交差検証R | 0.9337 | 最良自由度 | 3 |
| x3 | 3.3141 | 0.4227 | 1.1233 | | | | |
| x4 | 0.3820 | 0.5214 | | | | | |
| 切片 | 7.3511 | 0.0000 | | | | | |

図 21 PLS 回帰分析統計量

「最良自由度」は、目的変数の最適な自由度である。PLS 回帰分析の寄与率の値は、他に比べると少し落ちているが、リッジ回帰分析、PLS 回帰分析とも重回帰分析で問題となった変数 x2 と x4 の係数が小さくなっている。これにより、これらの変数のずれの影響は小さく抑えられる。

12.3 複数変数の異常検知

多変量正規分布に基づく異常検知

一般に p 変数の多変量正規分布の密度関数は以下で与えられる。

$$f(\mathbf{x} | \boldsymbol{\mu}, \boldsymbol{\Sigma}) = \frac{1}{\sqrt{(2\pi)^p |\boldsymbol{\Sigma}|}} \exp \left[-\frac{1}{2} {}^t(\mathbf{x} - \boldsymbol{\mu}) \boldsymbol{\Sigma}^{-1} (\mathbf{x} - \boldsymbol{\mu}) \right]$$

データ $D = \{\mathbf{x}_1, \dots, \mathbf{x}_\lambda\}$ が与えられた場合の対数尤度関数は以下で与えられる。

$$L(\boldsymbol{\mu}, \boldsymbol{\Sigma} | D) = -\frac{pN}{2} \log(2\pi) - \frac{N}{2} \log |\boldsymbol{\Sigma}| - \frac{1}{2} \sum_{\lambda=1}^N {}^t(\mathbf{x}_\lambda - \boldsymbol{\mu}) \boldsymbol{\Sigma}^{-1} (\mathbf{x}_\lambda - \boldsymbol{\mu})$$

我々は最尤法を用いて(2)式を最大化するが、その解は以下となる。

$$\hat{\boldsymbol{\mu}} = \frac{1}{N} \sum_{\lambda=1}^N \mathbf{x}_\lambda, \quad \hat{\boldsymbol{\Sigma}} = \frac{1}{N} \sum_{\lambda=1}^N (\mathbf{x}_\lambda - \hat{\boldsymbol{\mu}}) {}^t(\mathbf{x}_\lambda - \hat{\boldsymbol{\mu}})$$

ここで、同じ正規分布の確率変数 \mathbf{x}' に対する異常度 $a(\mathbf{x}')$ を $-2 \log f(\mathbf{x}' | \hat{\boldsymbol{\mu}}, \hat{\boldsymbol{\Sigma}})$ を元に以下のように定義する。

$$a(\mathbf{x}') = {}^t(\mathbf{x}' - \hat{\boldsymbol{\mu}}) \hat{\boldsymbol{\Sigma}}^{-1} (\mathbf{x}' - \hat{\boldsymbol{\mu}})$$

ここで上式と $-2 \log f(\mathbf{x}' | \hat{\boldsymbol{\mu}}, \hat{\boldsymbol{\Sigma}})$ の差は定数であるので、評価関数として本質的な差はない。

また上式は 1 次元変数の場合の変数の標準化の一般形である。

異常度の式については、以下のように定数を掛けると、分布が自由度 $p, N-p$ の F 分布に従うことが知られている。

$$T^2 \equiv \frac{N-p}{(N+1)p} {}^t(\mathbf{x}'_\lambda - \hat{\boldsymbol{\mu}}) \hat{\boldsymbol{\Sigma}}^{-1} (\mathbf{x}'_\lambda - \hat{\boldsymbol{\mu}}) \sim F_{p, N-p}$$

この T^2 をホテリング統計量という。

異常検知には、この統計量を使って確率の値を指定するか、直接 T^2 値を指定して閾値とする。

混合多変量正規分布に基づく異常検知

p 変数、 n 群混合多変量正規分布の密度関数は、群 α の確率密度関数

$$f_\alpha(\mathbf{x} | \boldsymbol{\mu}_\alpha, \boldsymbol{\Sigma}_\alpha) = \frac{1}{\sqrt{(2\pi)^p |\boldsymbol{\Sigma}_\alpha|}} \exp \left[-\frac{1}{2} {}^t(\mathbf{x} - \boldsymbol{\mu}_\alpha) \boldsymbol{\Sigma}_\alpha^{-1} (\mathbf{x} - \boldsymbol{\mu}_\alpha) \right] \quad (\alpha = 1, \dots, n)$$

を利用して以下で与えられる。

$$\begin{aligned} f(\mathbf{x} | \boldsymbol{\mu}_1, \dots, \boldsymbol{\mu}_n, \boldsymbol{\Sigma}_1, \dots, \boldsymbol{\Sigma}_n) &= \sum_{\alpha=1}^n \pi_\alpha f_\alpha(\mathbf{x} | \boldsymbol{\mu}_\alpha, \boldsymbol{\Sigma}_\alpha) \\ &= \sum_{\alpha=1}^n \frac{\pi_\alpha}{\sqrt{(2\pi)^p |\boldsymbol{\Sigma}_\alpha|}} \exp \left[-\frac{1}{2} {}^t(\mathbf{x} - \boldsymbol{\mu}_\alpha) \boldsymbol{\Sigma}_\alpha^{-1} (\mathbf{x} - \boldsymbol{\mu}_\alpha) \right] \end{aligned}$$

ここに、 π_α は群 α の生起確率である。

この密度関数に従うデータ $D = \{\mathbf{x}_1, \dots, \mathbf{x}_\lambda\}$ による対数尤度は以下である。

$$\begin{aligned} L(\pi_1, \dots, \pi_n, \boldsymbol{\mu}_1, \dots, \boldsymbol{\mu}_n, \boldsymbol{\Sigma}_1, \dots, \boldsymbol{\Sigma}_n | D) &= \sum_{\lambda=1}^N \log \left[\sum_{\alpha=1}^n \pi_\alpha f_\alpha(\boldsymbol{\mu}_\alpha, \boldsymbol{\Sigma}_\alpha | \mathbf{x}_\lambda) \right] \\ &= \sum_{\lambda=1}^N \log \left[\sum_{\alpha=1}^n \frac{\pi_\alpha}{\sqrt{(2\pi)^p |\boldsymbol{\Sigma}_\alpha|}} \exp \left\{ -\frac{1}{2} {}^t(\mathbf{x}_\lambda - \boldsymbol{\mu}_\alpha) \boldsymbol{\Sigma}_\alpha^{-1} (\mathbf{x}_\lambda - \boldsymbol{\mu}_\alpha) \right\} \right] \end{aligned}$$

最尤法を用いてこの対数尤度の最大値を求めるが、その際以下のアルゴリズムを利用する。

- 1) パラメータ $\pi_\alpha, \boldsymbol{\mu}_\alpha, \boldsymbol{\Sigma}_\alpha$ に初期値 $\hat{\pi}_\alpha, \hat{\boldsymbol{\mu}}_\alpha, \hat{\boldsymbol{\Sigma}}_\alpha$ を与える。
- 2) $\hat{\pi}_\alpha, \hat{\boldsymbol{\mu}}_\alpha, \hat{\boldsymbol{\Sigma}}_\alpha$ の値を用いて、各データの群 α への帰属度 $q_\alpha(\mathbf{x}_\lambda)$ を以下で求める。

$$q_\alpha(\mathbf{x}_\lambda) = \frac{\hat{\pi}_\alpha f_\alpha(\mathbf{x}_\lambda | \hat{\boldsymbol{\mu}}_\alpha, \hat{\boldsymbol{\Sigma}}_\alpha)}{\sum_{\beta=1}^n \hat{\pi}_\beta f_\beta(\mathbf{x}_\lambda | \hat{\boldsymbol{\mu}}_\beta, \hat{\boldsymbol{\Sigma}}_\beta)}$$

- 3) この帰属度を使い、新しいパラメータを以下のように決定する。

$$\begin{aligned} \hat{\pi}_\alpha &= \frac{1}{N} \sum_{\lambda=1}^N q_\alpha(\mathbf{x}_\lambda), \quad \hat{\boldsymbol{\mu}}_\alpha = \sum_{\lambda=1}^N q_\alpha(\mathbf{x}_\lambda) \mathbf{x}_\lambda / \sum_{\lambda=1}^N q_\alpha(\mathbf{x}_\lambda), \\ \hat{\boldsymbol{\Sigma}}_\alpha &= \sum_{\lambda=1}^N q_\alpha(\mathbf{x}_\lambda) (\mathbf{x}_\lambda - \hat{\boldsymbol{\mu}}_\alpha) {}^t(\mathbf{x}_\lambda - \hat{\boldsymbol{\mu}}_\alpha) / \sum_{\lambda=1}^N q_\alpha(\mathbf{x}_\lambda) \end{aligned}$$

- 4) 新しいパラメータと元のパラメータを比較し、十分近ければ（プログラムではすべての成分が 0.001 未満）終了し、そうでなければ 2) へ戻る。

この方法によって求めたパラメータを使って、異常判定には以下の指標を用いる。

$$a(\mathbf{x}') = -\log \sum_{\alpha=1}^n \hat{\pi}_{\alpha} f_{\alpha}(\mathbf{x}' | \hat{\boldsymbol{\mu}}_{\alpha}, \hat{\boldsymbol{\Sigma}}_{\alpha})$$

判定基準はこの指標を小さい方から順番に並べ、分位点を閾値として決めるか、直接指標の閾値を指定する。

このモデルの適合度は赤池情報量基準 AIC、ベイズ情報量基準 BIC などを使って求める。今このモデルのパラメータ数を M_n とすると、AIC と BIC はそれぞれ以下のように表現される。

$$AIC = -2L(\Theta | D) + 2M_n$$

$$BIC = -2L(\Theta | D) + M_n \log N, \quad M_n = \frac{p}{2}(n+1)(n+2)$$

ここに、 $L(\Theta | D)$ はパラメータを Θ で代表させて書いた対数尤度である。具体的には(8)式に求めたパラメータの値を代入したものである。適合度を求めるために、交差検証を使う方法も考えられるが、プログラムでは使用していない。

12.4 時系列データの異常検知

時系列データの異常検知で、周波数変化や波形変化に対応した検知法に特異スペクトル変換法がある。これは、時系列データ x_{t-w} から x_{t-1} までの w 個のデータを抽出し、それを 1 つのベクトル ${}^t\mathbf{x}_{t-w} = (x_{t-w} \ x_{t-w+1} \ \cdots \ x_{t-1})$ とする。そのベクトルの開始時点をずらして k ($k < w$) 本並べ、以下のような行列を作る。

$$\mathbf{X}_1^{(t)} = (\mathbf{x}_{t-k-w+1} \ \mathbf{x}_{t-k-w+2} \ \cdots \ \mathbf{x}_{t-w})$$

これに対して時間 L だけ経過した時点から作ったデータを以下のように $\mathbf{X}_2^{(t)}$ とする。

$$\mathbf{X}_2^{(t)} = (\mathbf{x}_{t-k-w+1+L} \ \mathbf{x}_{t-k-w+2+L} \ \cdots \ \mathbf{x}_{t-w+L})$$

この二つのデータの間で違いを見ることになる。

今、 ${}^t\mathbf{X}_1^{(t)}$, ${}^t\mathbf{X}_2^{(t)}$ の各列ベクトルに係数をかけて足し合わせ、特徴的な量を求めるために、その大きさを以下のように最大化する。

$$\|{}^t\mathbf{X}_1^{(t)}\mathbf{u}^{(t)}\|^2 \rightarrow \text{最大化} \quad \text{但し、} {}^t\mathbf{u}^{(t)}\mathbf{u}^{(t)} = 1, \quad \mathbf{u}^{(t)} \text{ は } (w \times 1) \text{ ベクトル}$$

$$\|{}^t\mathbf{X}_2^{(t)}\mathbf{v}^{(t)}\|^2 \rightarrow \text{最大化} \quad \text{但し、} {}^t\mathbf{v}^{(t)}\mathbf{v}^{(t)} = 1, \quad \mathbf{v}^{(t)} \text{ は } (w \times 1) \text{ ベクトル}$$

このベクトルは、以下の固有方程式から求められ、

$$\mathbf{X}_1^{(t)} {}^t\mathbf{X}_1^{(t)} \mathbf{u}^{(t)} = \lambda \mathbf{u}^{(t)}$$

$${}^t\mathbf{X}_2^{(t)} {}^t\mathbf{X}_2^{(t)} \mathbf{v}^{(t)} = \mu \mathbf{v}^{(t)}$$

固有値の大きい順に固有ベクトルを m 個並べて表した $(w \times m)$ 行列を以下のように定義す

る。

$$\mathbf{U}_m^{(t)} (w \times m) = (\mathbf{u}_1^{(t)} \mathbf{u}_2^{(t)} \cdots \mathbf{u}_m^{(t)})$$

$$\mathbf{V}_m^{(t)} (w \times m) = (\mathbf{v}_1^{(t)} \mathbf{v}_2^{(t)} \cdots \mathbf{v}_m^{(t)})$$

これら 2 つの行列を用いて、変化度 $a(t)$ は行列 ${}^t\mathbf{U}_m^{(t)}\mathbf{V}_m^{(t)} (m \times m)$ の 2 ノルムを用いて以下のように定義される。

$$a(t) = 1 - \left\| {}^t\mathbf{U}_m^{(t)}\mathbf{V}_m^{(t)} \right\|_2^2 = 1 - \left({}^t\mathbf{U}_m^{(t)}\mathbf{V}_m^{(t)} \text{ の最大特異値} \right)^2$$

ここに、実行列 \mathbf{A} の 2 ノルム $\|\mathbf{A}\|_2$ は以下のように定義される。

$$\|\mathbf{A}\|_2 = \max_{\mathbf{u}} \sqrt{\frac{{}^t(\mathbf{A}\mathbf{u})\mathbf{A}\mathbf{u}}{{}^t\mathbf{u}\mathbf{u}}} = \sqrt{\lambda_{\max}}$$

注) 行列 \mathbf{A} の特異値とは、行列 \mathbf{A} と \mathbf{A} の随伴行列 \mathbf{A}^* の積 $\mathbf{A}\mathbf{A}^*$ ($\mathbf{A}^*\mathbf{A}$) の非負の固有値の平方根である。

12.5 入力と出力がある異常検知

入力と出力の関係で生じる異常の検知方法については、重回帰分析を用いる手法が考えられる。しかし、重回帰分析は、入力変数が多くその値が似通っている場合に、多重共線性の問題が発生する可能性があり、予測が不安定となる。これに対して改善方法と考えられている代表的な手法がリッジ回帰分析と PLS 回帰分析である。リッジ回帰分析は、多重共線性の元となる分散共分散行列に手を加える手法であり、PLS 回帰分析は多重共線性を与える変数間の自由度を制約する手法である。我々のプログラムは 3 者を比較するように作成しており、その違いを理解し易くなっている。

重回帰分析

重回帰分析の目的変数を y_λ ($\lambda=1,2,\dots,N$)、説明変数を $x_{i\lambda}$ ($i=1,2,\dots,p$) とし、それらの関係を ε_λ を誤差項として以下とする。

$$y_\lambda = \sum_{i=1}^p b_i x_{i\lambda} + b_0 + \varepsilon_\lambda$$

最小 2 乗法としての重回帰分析では、以下の値 D が最小になるように、パラメータ b_i, b_0 を決定する。

$$D = \sum_{\lambda=1}^N \left(y_\lambda - \sum_{i=1}^p b_i x_{i\lambda} - b_0 \right)^2 = {}^t(\mathbf{y} - \mathbf{X}\mathbf{b})(\mathbf{y} - \mathbf{X}\mathbf{b})$$

ここに、

$$(\mathbf{X})_{i\lambda} = \tilde{x}_{i\lambda} = x_{i\lambda} - \bar{x}_i, \quad (\mathbf{y})_\lambda = \tilde{y}_\lambda = y_\lambda - \bar{y}, \quad \mathbf{b} = {}^t(b_1, b_2, \dots, b_p)$$

である。パラメータは以下で与えられる。

$$\mathbf{b} = ({}^t\mathbf{X}\mathbf{X})^{-1} {}^t\mathbf{X}\mathbf{y}, \quad b_0 = \bar{y} - \sum_{i=1}^p b_i \bar{x}_i$$

問題となる多重共線性は、行列 ${}^t\mathbf{XX}$ の非正則性から生じる。

多重共線性について、 i 番目の説明変数を、他の説明変数で予測して重相関係数 r_i を求め、以下の式で定義される VIF 指標を利用している。

$$VIF_i = 1/(1-r_i^2)$$

異常度について、通常の重回帰分析では以下で定義する。

$$a(y', \mathbf{x}') = \frac{(y' - b_0 - \mathbf{b}\mathbf{x}')^2}{\sigma^2}$$

ここで分散 σ^2 については以下で推測する。

$$\sigma^2 = \frac{D}{N} = \frac{1}{N} {}^t(\mathbf{y} - \mathbf{X}\mathbf{b})(\mathbf{y} - \mathbf{X}\mathbf{b})$$

リッジ回帰分析

リッジ回帰分析は重回帰分析の正則性の問題に対して、以下のように置くことによって正則性を確保しようとする手法である。

$$\mathbf{b}' = ({}^t\mathbf{XX} + \eta\mathbf{I})^{-1} {}^t\mathbf{X}\mathbf{y}$$

これは、以下を最小化する解でもある。

$$D' = {}^t(\mathbf{y} - \mathbf{X}\mathbf{b}')(\mathbf{y} - \mathbf{X}\mathbf{b}') + \eta {}^t\mathbf{b}'\mathbf{b}'$$

ここでパラメータ η の値は以下のようにして求められる。 α 番目の個体を抜いた 1 個抜き交差検証のパラメータを $\mathbf{b}^{(-\alpha)}$ とすると、そのときの平均 2 乗誤差 $e(\eta)$ は以下で与えられることが知られている。

$$e(\eta) = \frac{1}{N} \sum_{\lambda=1}^N (\tilde{y}_\lambda - \sum_{i=1}^p \tilde{x}_{i\lambda} b_i')^2$$

これは、以下のように書くこともできる。

$$e(\eta) = \frac{1}{N} {}^t\mathbf{A}\mathbf{A}$$

ここに、

$$\mathbf{A} = \text{diag}(\mathbf{I} - \mathbf{H})^{-1}(\mathbf{I} - \mathbf{H})\mathbf{y}, \quad \mathbf{H}(N \times N) = \mathbf{X}({}^t\mathbf{XX} + \eta\mathbf{I})^{-1} {}^t\mathbf{X}$$

また、 $\text{diag}(\mathbf{I} - \mathbf{H})^{-1}$ は対角要素が $(\mathbf{I} - (\mathbf{H})_{ii})^{-1}$ となる対角行列である。運用上はパラメータ η の値を変化させて、この $e(\eta)$ が最小になるようなパラメータ η を選ぶ。

もう少し安全性を考えて、以下の一般化交差確認検証法と呼ばれる方法から与えられる誤差 $e_{GCV}(\eta)$ を最小化する場合もある。

$$e_{GCV}(\eta) = \frac{1}{N} {}^t\mathbf{A}'\mathbf{A}'$$

ここに、 $\mathbf{A}' = (\mathbf{I} - \mathbf{H})\mathbf{y}/[1 - \text{tr}\mathbf{H}/N]$ である。我々のプログラムでは前者の判定法を利用している。

多重共線性がある場合、重回帰分析の予測は、そのデータに対してだけは良い精度を与えるが、他の新しいデータを用いた場合、予測の精度が著しく低下する。そのため 1 個抜き

交差検証は必須である。

異常度について、リッジ回帰分析では以下で定義する。

$$a(y', \mathbf{x}') = \frac{(y' - b'_0 - \mathbf{b}'\mathbf{x}')^2}{\sigma'^2}$$

ここで分散 σ'^2 については以下で推測する。

$$\sigma'^2 = \frac{D'}{N} = \frac{1}{N} [\mathbf{y}'(\mathbf{y} - \mathbf{X}\mathbf{b})(\mathbf{y} - \mathbf{X}\mathbf{b}) + \eta' \mathbf{b}'\mathbf{b}']$$

我々のプログラムでは、重相関係数と寄与率について、残差分散を以下で計算している。

$$\sigma^2 = \frac{1}{N} \mathbf{y}'(\mathbf{y} - \mathbf{X}\mathbf{b})(\mathbf{y} - \mathbf{X}\mathbf{b})$$

そのため、出力される分散の値が他の分析（例えば以下の PLS 回帰）より大きいのに、重相関係数の値が大きく表示されるということもある。

PLS 回帰分析

PLS 回帰分析ではまず、変数の線形結合を考える。

$$r_{i\lambda} = \sum_{j=1}^p u_{ij} \tilde{x}_{j\lambda} \quad (i=1, 2, \dots, r; r < p)$$

この式を行列記号を用いて書くと以下となる。

$$\mathbf{R} = \mathbf{X}\mathbf{U} \quad \mathbf{U} = (\mathbf{u}_1 \mathbf{u}_2 \cdots \mathbf{u}_r)$$

ここで、行列 \mathbf{U} の各列ベクトルは直交し、順番に $\mathbf{X}\mathbf{u}_i$ と \mathbf{y} との内積が最大化されるように選ばれる。詳細は後に示す。

この新しい変数を用いて、目的変数を以下のように予測する。

$$\tilde{y}_\lambda = \sum_{j=1}^r \beta_j r_{j\lambda} + \varepsilon_\lambda$$

即ち、

$$\mathbf{y} = \mathbf{R}\boldsymbol{\beta} + \boldsymbol{\varepsilon} = \mathbf{X}\mathbf{U}\boldsymbol{\beta} + \boldsymbol{\varepsilon}$$

最小 2 乗法を使うと、以下の量を最小化するようにパラメータを決定する。

$$D'' = \mathbf{y}'(\mathbf{y} - \mathbf{R}\boldsymbol{\beta})(\mathbf{y} - \mathbf{R}\boldsymbol{\beta})$$

その解は次のように与えられる。

$$\boldsymbol{\beta} = (\mathbf{R}'\mathbf{R})^{-1} \mathbf{R}'\mathbf{y} = (\mathbf{U}'\mathbf{X}\mathbf{X}\mathbf{U})^{-1} \mathbf{U}'\mathbf{X}\mathbf{y}$$

これから、標準化偏回帰係数 $\tilde{\mathbf{b}}$ は以下となる。

$$\tilde{\mathbf{b}} = \mathbf{U}\boldsymbol{\beta}$$

また、通常の回帰係数は、以下で与えられる。

$$b_i'' = \tilde{b}_i s_{y_i} / s_i, \quad b_0'' = \bar{y} - \sum_{i=1}^p b_i'' \bar{x}_i$$

多重共線性の改善の程度については、変数を \mathbf{U} 行列で変換した後の i 番目の説明変数を、他の説明変数で予測して重相関係数 r_i を求め、以下の式で定義される VIF 指標を利用して

いる。

$$VIF_i = 1/(1-r_i^2)$$

異常度については以下で定義する。

$$a(y', \mathbf{x}') = \frac{(y' - b_0'' - \mathbf{b}''\mathbf{x}')^2}{\sigma^2}$$

ここで分散 σ^2 については以下で推測する。

$$\sigma^2 = \frac{D}{N} = \frac{1}{N} {}^t(\mathbf{y} - \mathbf{X}\mathbf{b}'')(\mathbf{y} - \mathbf{X}\mathbf{b}'')$$

最後に行列 $\mathbf{U} = (\mathbf{u}_1 \mathbf{u}_2 \cdots \mathbf{u}_r)$ の決定法について述べる。この行列の 1 列目 \mathbf{u}_1 は $\mathbf{X}\mathbf{u}_1$ が最も \mathbf{y} の方向に向くように、以下のように求める。

$$L_1 = {}^t\mathbf{y}\mathbf{X}\mathbf{u}_1 - \mu_1({}^t\mathbf{u}_1\mathbf{u}_1 - 1) \rightarrow \text{最大化}$$

この解は以下で与えられる。

$$\mathbf{u}_1 = {}^t\mathbf{X}\mathbf{y} / \|{}^t\mathbf{X}\mathbf{y}\|$$

次の \mathbf{u}_2 については、 \mathbf{X} から \mathbf{u}_1 方向の成分を取り除き、以下のように求める。

$$L_2 = {}^t\mathbf{y}(\mathbf{X} - \mathbf{d}_1 {}^t\mathbf{d}_1\mathbf{X})\mathbf{u}_2 - \mu_2({}^t\mathbf{u}_2\mathbf{u}_2 - 1) \rightarrow \text{最大化}$$

ここに、 $\mathbf{d}_1 = \mathbf{X}\mathbf{u}_1 / \|\mathbf{X}\mathbf{u}_1\|$ である。確かに $\mathbf{X} - \mathbf{d}_1 {}^t\mathbf{d}_1\mathbf{X}$ は、以下のように \mathbf{u}_1 方向の成分を取り除いている。

$$(\mathbf{X} - \mathbf{d}_1 {}^t\mathbf{d}_1\mathbf{X})\mathbf{u}_1 = \mathbf{X}\mathbf{u}_1 - \mathbf{X}\mathbf{u}_1 {}^t(\mathbf{X}\mathbf{u}_1)\mathbf{X}\mathbf{u}_1 / \|\mathbf{X}\mathbf{u}_1\|^2 = \mathbf{0}$$

この解は以下で与えられる。

$$\mathbf{u}_2 = {}^t(\mathbf{X} - \mathbf{d}_1 {}^t\mathbf{d}_1\mathbf{X})\mathbf{y} / \|{}^t(\mathbf{X} - \mathbf{d}_1 {}^t\mathbf{d}_1\mathbf{X})\mathbf{y}\|$$

このベクトル \mathbf{u}_2 は以下のように \mathbf{u}_1 と直交する。

$${}^t\mathbf{u}_1\mathbf{u}_2 \propto {}^t\mathbf{u}_1 {}^t(\mathbf{X} - \mathbf{d}_1 {}^t\mathbf{d}_1\mathbf{X})\mathbf{y} = {}^t(\mathbf{X}\mathbf{u}_1 - \mathbf{X}\mathbf{u}_1 {}^t(\mathbf{X}\mathbf{u}_1)\mathbf{X}\mathbf{u}_1 / \|\mathbf{X}\mathbf{u}_1\|^2)\mathbf{y} = 0$$

これを続けると、 k 番目の係数ベクトル \mathbf{u}_k は以下のように求められることが分かる。

$$L_k = {}^t\mathbf{y}\left(\mathbf{X} - \sum_{i=1}^{k-1} \mathbf{d}_i {}^t\mathbf{d}_i\mathbf{X}\right)\mathbf{u}_k - \mu_k({}^t\mathbf{u}_k\mathbf{u}_k - 1) \rightarrow \text{最大化}$$

$$\mathbf{u}_k = {}^t\left(\mathbf{X} - \sum_{i=1}^{k-1} \mathbf{d}_i {}^t\mathbf{d}_i\mathbf{X}\right)\mathbf{y} / \left\|{}^t\left(\mathbf{X} - \sum_{i=1}^{k-1} \mathbf{d}_i {}^t\mathbf{d}_i\mathbf{X}\right)\mathbf{y}\right\|$$

どこまでの次元数を求めればよいかは、1つの方法として1個抜き交差検証法の重相関係数または残差分散の大きさを元にして決めればよい。我々のプログラムではこの方法を用いている。

参考文献

- [1] 井出剛, 入門機械学習による異常検知, コロナ社, 2015.

1.3. ナッシュ均衡ツール

経営分野でよく利用される OR の手法にゲーム理論がある。ゲーム理論は、大きく分けてゼロ和ゲームと非ゼロ和ゲームに分類され、非ゼロ和ゲームの中に非協力ゲームと協力ゲームがある。今回プログラム化したものは、ゼロ和 2 人ゲームと非ゼロ和 2 人非協力ゲームである。それぞれ 1 回の打ち手で均衡解が得られる純粋戦略の解と、1 回では均衡解が得られず、無限回の確率的な打ち手で均衡解が得られる混合戦略の解がある。ここで述べる均衡解とは、お互い相手の戦略が変化しないとした場合、自分の戦略を変える動機が生じない場合をいう。すなわち相手の戦略を所与とした場合、お互い最良の選択をしている場合のことをいう。

ここではゲームの利得が有限な行列の形で与えられる問題を考える。ゼロ和ゲームの場合は 1 つの行列でプレイヤー 1 の利得とプレイヤー 2 の損失を表すことが多く、非ゼロ和ゲームの場合は 2 つの行列で、それぞれのプレイヤーの利得を表すことが多い。ここではそれぞれのプレイヤーについて利得行列か損失行列かを選択できるようにしている。

このプログラムでは解を 3 つの方法で求めている。1 つ目は純粋戦略の方法、2 つ目は微分を使った方法、3 つ目は線形計画法である。微分を使った方法については補遺で述べる。また、線形計画法は混合戦略ゼロ和 2 人ゲームの解法である。

最初に、最もよく知られた 2×2 の利得行列の問題を使ってプログラムの簡単な使い方を説明し、その後、一般の行列の 2 人ゼロ和ゲーム、非協力 2 人非ゼロ和ゲームの場合に進む。

1. 2×2 行列のゲーム

C.Analysis のメニュー [分析－OR－ナッシュ均衡ツール] を選択すると、図 1 に示す分析実行画面が表示される。

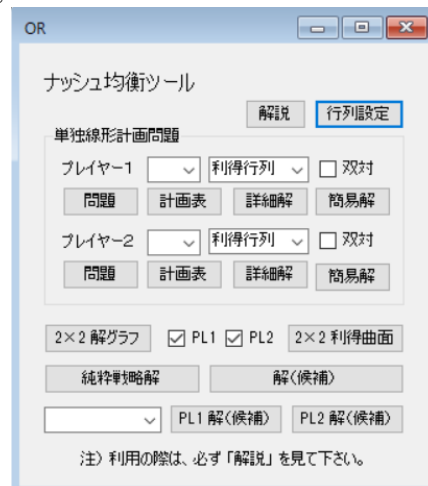


図 1 分析実行画面

データの利得・損失行列は図 2 のように、C.Analysis のメニュー [分析－数学－行列計算] の書式で表す。ここではまず 2×2 行列の例を示す。

| データ編集 ナッシュ均衡1.txt | | | | | | | |
|-------------------|-------|-----|----|--|----|----|----|
| <div> </div> | | | | | | | |
| ▶ | a= | 4 | 0 | | b= | 4 | 6 |
| | p70 | 6 | 2 | | | 0 | 2 |
| | c= | -24 | 4 | | d= | 12 | -6 |
| | p70 | 6 | -2 | | | -8 | 2 |
| | e= | 0 | 10 | | f= | 0 | 6 |
| | p71 | 6 | -6 | | | 10 | -6 |
| | g= | 2 | -1 | | h= | 1 | -1 |
| | p73 | -1 | 1 | | | -1 | 2 |
| | i= | 4 | 0 | | j= | 4 | 6 |
| | p75弱虫 | 6 | -2 | | | 0 | -2 |
| | | | | | | | |
| 2/3 (1,1) | | | | | | 備考 | |

図2 ナッシュ均衡のデータ形式（ナッシュ均衡 1.txt, 2 頁目）

行列は左上に「行列名=」を付けて、その右から下にかけて具体的な行列データを入力する。行列と行列の間は、1行1列以上空ける必要がある。また、行列名の下部分は検索しないので、コメントを加えることができる。

プログラムを動作させるには、まず分析実行画面右上の「行列設定」ボタンをクリックする。その際、図3のように、「単独線形計画問題」グループボックス内の「プレイヤー1」と「プレイヤー2」ラベルの横に行列の候補が表示され、その右のコンボボックスで利得行列か損失行列かが選択できるようになる。

図3 利得・損失行列の設定

例1 プレイヤー1：利得行列 a，プレイヤー2：利得行列 b

図2において、プレイヤー1の利得行列をa、プレイヤー2の利得行列をbとして、「解（候補）」のボタンをクリックしてみる。ここで、「解（候補）」となっているのは、戦略数が多いときには完全な解を表示するとは限らないからであるが、これについては後に述べる。この問題は囚人のジレンマとしてよく知られた問題である。結果を図4に示す。

| 戦路解 (...) | |
|-----------|-------|
| 純粋戦略1 | |
| ▶ プレイヤー1 | |
| p1 | 0.000 |
| p2 | 1.000 |
| 最大化解 | 2.000 |
| プレイヤー2 | |
| q1 | 0.000 |
| q2 | 1.000 |
| 最大化解 | 2.000 |

図4 解（候補）

これは、プレイヤー1の戦略1の選択確率 p_1 が0、戦略2の選択確率 p_2 が1で最大利得が2、プレイヤー2の戦略1の選択確率 q_1 が0、戦略2の選択確率 q_2 が1で最大利得が2であることを示しており、純粋戦略の解である。

この解を別の形で表してみる。分析実行画面の「純粋戦略解」をクリックすると、図5のような結果が得られる。

| 純粋戦略解 (A, B) | | |
|--------------|--------|----------|
| 利得\利得 | PL2戦略1 | PL2戦略2 |
| ▶ PL1戦略1 | 4, 4 | 0, (6) |
| PL1戦略2 | (6), 0 | (2), (2) |

図5 純粋戦略解

これは練習問題などでよく見られる形式で、相手の戦略に応じた最適解の利得が「(6)」などのように、選択される可能性のある利得が括弧で囲まれて表示される。またプレイヤー1の最適解は青色、プレイヤー2の最適解は緑色、双方の最適解、即ち均衡解は黄色で示されている。

同じくもう1つの表し方を見てみよう。分析実行画面の「2×2 解グラフ」ボタンをクリックすると、図6のようなグラフが表示される。



図6 2×2 解グラフ

これは横軸が p_1 、縦軸が q_1 を表し、青い線と緑の線の交わる場所 $p_1=0, q_1=0$ が均衡解である。図4で見たように、解は $p_1=0, q_1=0$ であるが、図5の戦略で見ると、双方とも戦略2が最善戦略であり、同じ結果を示している。この線の引き方については、よく知られているが、補遺1に改めて説明しておく。

次にプレイヤー1とプレイヤー2の利得（損失）曲面を表示してみよう。分析実行画面の「2×2 利得曲面」ボタンをクリックすると図7左のような「2変数関数グラフ」の実行画

面が表示され、そのまま「グラフ描画」ボタンをクリックすると図 7 右のような 3 次元グラフが表示される。ここで、グラフの色は「標準色」チェックボックスのチェックを外し、後で設定したものである。また、赤い丸は場所を分かり易くするために後から付け加えたものである。

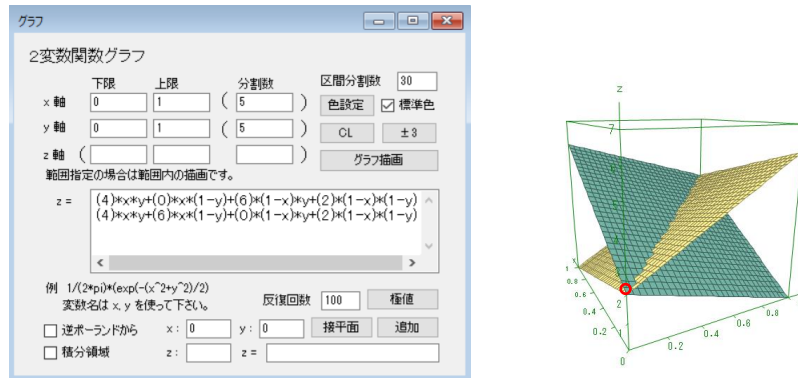


図 7 2×2 利得曲面

解は、双方とも最良の選択であるので、プレイヤー 1 では x 軸方向に解を移動させると利得は下がる解として $x=0$ が選ばれ、プレイヤー 2 では y 軸方向に解を移動させると利得が下がる解として $y=0$ が選ばれる。これが図 6 の 2 本の直線であり、交点が均衡解である。

例 2 プレイヤー 1 : 利得行列 c , プレイヤー 2 : 利得行列 d

| 戦略解 (C, D) | |
|------------|--------|
| プレイヤー 1 | 混合戦略 1 |
| p1 | 0.357 |
| p2 | 0.643 |
| 最大化解 | -0.667 |
| プレイヤー 2 | |
| q1 | 0.167 |
| q2 | 0.833 |
| 最大化解 | -0.857 |

| 利得\利得 | PL2戦略1 | PL2戦略2 |
|--------|-----------|---------|
| PL1戦略1 | -24, (12) | (4), -6 |
| PL1戦略2 | (6), -8 | -2, (2) |

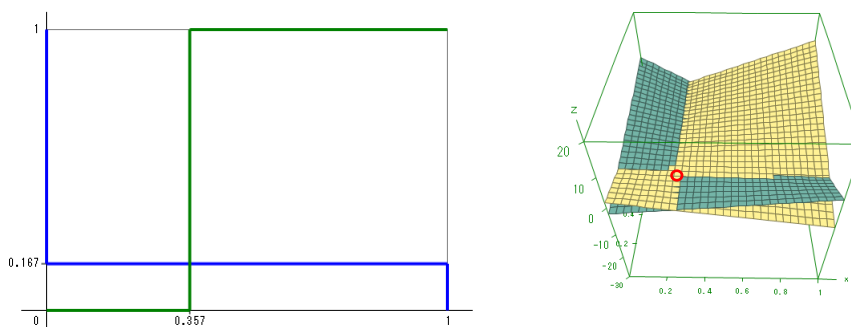


図 8 混合戦略の解だけを持つ例

これは混合戦略の解だけを持つ例であり、純粋戦略の解は存在しない。それぞれのプレイヤーは相手が値を変更しない限り、自分が値を変更しても利得は不変である。この解は線形計画法を使っても求めることができる。2つのゼロ和2人ゲームの双対価格が解である。

例3 プレイヤー1：利得行列 e, プレイヤー2：利得行列 f

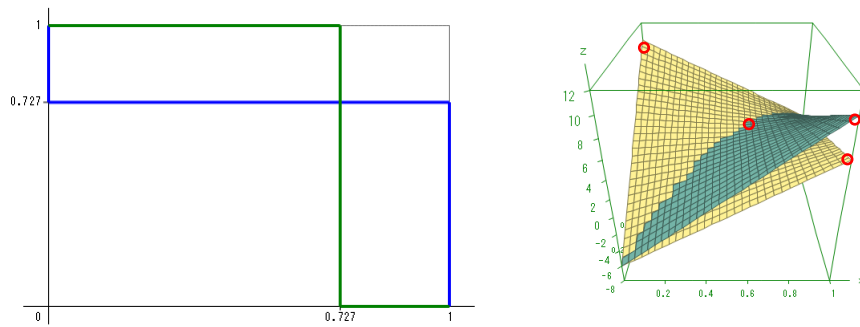


図9 純粋戦略と混合戦略の解を持つ例1

これは純粋戦略の解と混合戦略の解、両方を持つ例である。この解は、線形計画法を使って求めることができる。

例4 プレイヤー1：利得行列 i, プレイヤー2：利得行列 j

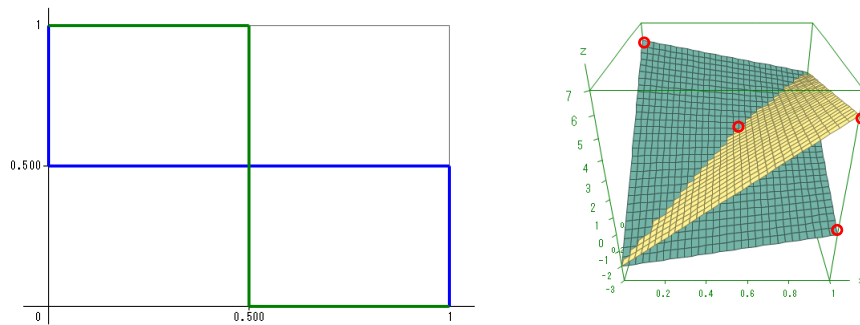


図10 純粋戦略と混合戦略の解を持つ例2

これも純粋戦略の解と混合戦略の解、両方を持つ例である。この解は、線形計画法を使っては求めることができない。

2. 一般の行列のゼロ和2人ゲーム

一般の利得・損失行列の場合は図1の行列の例を用いて説明する。

| a= | | | | b= | | | |
|-----|-----|----|--|----|----|---|----|
| 4 | -1 | -2 | | 5 | 2 | 1 | |
| -3 | 0 | 3 | | 3 | 2 | 2 | |
| 3 | 1 | 2 | | 4 | 1 | 0 | |
| c= | | | | d= | | | |
| 0 | 150 | | | 5 | 2 | 4 | -1 |
| 100 | 0 | | | -2 | -1 | 2 | 3 |
| e= | | | | f= | | | |
| 5 | 1 | | | 5 | 5 | 2 | 1 |
| 0 | 3 | | | 3 | 1 | 2 | 4 |
| 2 | 4 | | | | | | |
| 3 | 0 | | | | | | |
| g= | | | | h= | | | |
| 5 | 2 | -2 | | 1 | 3 | 1 | |
| 2 | 1 | 3 | | 1 | 1 | 3 | |
| -1 | 3 | 6 | | 4 | 2 | 1 | |
| i= | | | | | | | |
| 1 | 2 | 3 | | | | | |
| 4 | 2 | 6 | | | | | |
| 5 | 2 | 2 | | | | | |

図1 ナッシュ均衡のデータ形式（ナッシュ均衡1.txt, 1 頁目）

ゼロ和2人ゲームの純粋戦略の解は、マクシミン戦略とミニマックス戦略を用いて与えられることが知られているが、プログラムではもう少し簡単に相手の戦略に対して最良の戦略を取るという考えで（双対問題として）求めている。また混合戦略の解は微分を使う方法で求めている。但し、これらの解は単一ではないので注意が必要である。

ここで、プレイヤー1の利得行列を a として、純粋戦略ゼロ和2人ゲームの問題を解いてみよう。まず、図2のように、プレイヤー1の行列を a として利得行列に、プレイヤー2の行列を同じく a として損失行列に設定する。

図2 実行画面の設定

その後「純粋戦略解」のボタンをクリックすると、図3のような解が表示される。

| 利得\損失 | PL2戦略1 | PL2戦略2 | PL2戦略3 |
|--------|----------|----------|----------|
| PL1戦略1 | (4), 4 | -1, -1 | -2, (-2) |
| PL1戦略2 | -3, (-3) | 0, 0 | (3), 3 |
| PL1戦略3 | 3, 3 | (1), (1) | 2, 2 |

図3 純粋戦略解

これは、相手が戦略 1, 2, 3 を取る場合、自分にとってどの戦略が最適かということを考えて求めた図である。結果の見方は 2×2 行列のときと同じである。黄色の部分の純粋戦略の解であるが、一般に 1 か所になるとは限らない。

次に線形計画法を使ってこの問題を解いてみよう。まず、プレイヤー 1 の行列を a にし、利得行列に設定する。設定はこれだけでよい。プレイヤー 1 の「問題」ボタンをクリックすると、図 4 のような線形計画問題の式が表示される。

```

z=u MAX
4*p1-3*p2+3*p3-u>=0
-p1+p3-u>=0
-2*p1+3*p2+2*p3-u>=0
p1+p2+p3=1
p1,p2,p3>=0

```

図4 線形計画問題

これは C.Analysis の線形計画法 (OR) のテキストエディタでの入力形式で、以下のよう
な問題を表したものである。

目的関数 $z = u$ 最大化

制約条件 $4p_1 - 3p_2 + 3p_3 \geq u$

$-p_1 + p_3 \geq u$

$-2p_1 + 3p_2 + 2p_3 \geq u$

$p_1 + p_2 + p_3 = 1$

$p_1, p_2, p_3 \geq 0$

この問題は、プレイヤー 2 が 3 つの戦略を選択したとき、プレイヤー 1 の利得の期待値がある数 u より大きいとして、その数 u を最大化する問題である。これを線形計画表として表示するには、「計画表」ボタンをクリックする。結果は図 5 のようになる。

| | p1 | p2 | p3 | u | |
|---|----|----|----|----|----------|
| ▶ | | | | 1 | MAX |
| | 4 | -3 | 3 | -1 | ≥ 0 |
| | -1 | 0 | 1 | -1 | ≥ 0 |
| | -2 | 3 | 2 | -1 | ≥ 0 |
| | 1 | 1 | 1 | | $= 1$ |

図5 線形計画表

これは、C.Analysis の線形計画法 (OR) のグリッドエディタでの入力形式である。この解は、「詳細解」をクリックすると図 6 のように与えられる。

| LP 最適解 | | STEP数 | 最大値(Z) |
|--------|-------|-------|--------|
| | | 7 | 1.000 |
| 変数 | 値 | 被約費用 | |
| p1 | 0.000 | 2.000 | |
| p2 | 0.000 | 1.000 | |
| p3 | 1.000 | 0.000 | |
| u_1 | 1.000 | 0.000 | |
| u_2 | 0.000 | 0.000 | |
| 行 | スラック | 双対価格 | |
| 1 | 2.000 | 0.000 | |
| 2 | 0.000 | 1.000 | |
| 3 | 1.000 | 0.000 | |
| 4 | 0.000 | 1.000 | |
| 係数範囲 | | 相対値 | 最小値 |

図 6 線形計画問題の詳細解

ここでは、解（プレイヤー 1 の確率）は黄色、双対価格（プレイヤー 2 の確率）はオレンジ色、目的関数値（ゲームの値または最終的な期待利得）はピンクに色付けられている。双対価格がプレイヤー 2 の確率になるのは、双対問題がプレイヤー 2 から見た損失行列の最小化問題になるからである。実際、分析実行画面でプレイヤー 1 の右端の「双対」チェックボックスにチェックを入れ、「問題」ボタンをクリックした結果が図 7 である。

```

z=v MIN
4*q1-q2-2*q3-v<=0
-3*q1+3*q3-v<=0
3*q1+q2+2*q3-v<=0
q1+q2+q3=1
q1,q2,q3>=0

```

図 7 プレイヤー 1 の線形計画問題の双対問題

この双対問題はプレイヤー 2 の損失の最小化問題である。

図 6 の結果表示は線形計画問題としては十分な表示であるが、均衡問題の解の表示法としては見易いとはいえない。そのため均衡問題に適した表示法も用意してある。「簡易解」ボタンをクリックすると、図 8 のような形式の解が得られる。

| Player 1 | | q1 | q2 | q3 | |
|----------|-------|--------|--------|--------|--------|
| 利得行列 | 確率 | 0.000 | 1.000 | 0.000 | |
| p1 | 0.000 | 4.000 | -1.000 | -2.000 | -1.000 |
| p2 | 0.000 | -3.000 | 0.000 | 3.000 | 0.000 |
| p3 | 1.000 | 3.000 | 1.000 | 2.000 | 1.000 |
| | | 3.000 | 1.000 | 2.000 | 1.000 |

図 8 均衡問題の簡易解

黄色、オレンジ色、ピンクは詳細解に準じている。右端の青色の部分は、プレイヤー 1 がその戦略を選んだ場合の利得の期待値を表している。そのため、プレイヤー 1 は利得の期待値が最も大きい戦略 3 を確率 1 で選んでいるということになる。逆にこれを縦に見ると、相手も損失の期待値が最も小さい戦略 2 を確率 1 で選んでいることになる。

次に混合戦略の問題を考えてみよう。行列 d を利得行列として選んだ簡易解は図 9 のようになる。

| Player 1 | | q1 | q2 | q3 | q4 | q5 | |
|----------|-------|--------|--------|-------|--------|-------|-------|
| ▶ 利得行列 | 確率 | 0.000 | 0.571 | 0.000 | 0.429 | 0.000 | |
| p1 | 0.571 | 5.000 | 2.000 | 4.000 | -1.000 | 0.000 | 0.714 |
| p2 | 0.429 | -2.000 | -1.000 | 2.000 | 3.000 | 3.000 | 0.714 |
| | | 2.000 | 0.714 | 3.143 | 0.714 | 1.286 | 0.714 |

図 9 利得行列 d の混合戦略解

これを見ると、プレイヤー 1 の 2 つの戦略に対して、利得の期待値は同じ 0.714 の値をとっている。これは確率を変えても期待値が変わらないことを示しており、自分の解を変える動機がないという均衡の考え方をよく表している。また、プレイヤー 2 にとっても戦略 2 と戦略 4 は最も損失の小さい同じ 0.714 の値をとっており、解を変える動機が存在しない。

混合戦略の解も唯一とは限らない。その例を示しておく。行列 f を利得行列として選んだ線形計画問題の簡易解を図 10a に与え、その双対問題の簡易解を図 10b に与える。

| Player 1 | | q1 | q2 | q3 | q4 | |
|----------|-------|-------|-------|-------|-------|-------|
| ▶ 利得行列 | 確率 | 0.000 | 0.000 | 1.000 | 0.000 | |
| p1 | 0.667 | 5.000 | 5.000 | 2.000 | 1.000 | 2.000 |
| p2 | 0.333 | 3.000 | 1.000 | 2.000 | 4.000 | 2.000 |
| | | 4.333 | 3.667 | 2.000 | 2.000 | 2.000 |

図 10a 利得行列 f の解

| Player 1 | | q1 | q2 | q3 | q4 | |
|----------|-------|-------|-------|-------|-------|-------|
| ▶ 利得行列 | 確率 | 0.000 | 0.000 | 1.000 | 0.000 | |
| p1 | 0.250 | 5.000 | 5.000 | 2.000 | 1.000 | 2.000 |
| p2 | 0.750 | 3.000 | 1.000 | 2.000 | 4.000 | 2.000 |
| | | 3.500 | 2.000 | 2.000 | 3.250 | 2.000 |

図 10b 利得行列 f の双対問題を解いた解

通常、主問題の解は双対問題の双対価格になっているが、この場合結果が異なっている。これは解が 1 つではなく、2 つの解を結ぶ線分上の値も解であるためである。

$$p_1 = 0.667 \times \alpha + 0.250 \times (1 - \alpha) \quad (0 \leq \alpha \leq 1)$$

$$p_2 = 0.333 \times \alpha + 0.750 \times (1 - \alpha)$$

このプログラムでは解が網羅できるわけではないことに気を付けてもらいたい。

次にこれを 2 つの行列に分けてやってみよう。プレイヤー 1 は利得行列として f、プレイヤー 2 は損失行列として f をとり、「解 (候補)」ボタンをクリックすると図 11 のような結果を得る。

| | 混合戦略 1 | 混合戦略 2 |
|-----------|--------|--------|
| ▶ プレイヤー 1 | | |
| p1 | 0.667 | 0.250 |
| p2 | 0.333 | 0.750 |
| 最大化解 | 2.000 | 2.000 |
| プレイヤー 2 | | |
| q1 | 0.000 | 0.000 |
| q2 | 0.000 | 0.000 |
| q3 | 1.000 | 1.000 |
| q4 | 0.000 | 0.000 |
| 最小化解 | 2.000 | 2.000 |

図 11 微分の方法で求めた解

ここでも両端の解だけが求められている。

3. 一般の行列の非ゼロ和 2 人非協力ゲーム

非ゼロ和ゲームの場合、プレイヤー 1 とプレイヤー 2 で同じ行数と列数の行列を 2 つ選択する。これらに利得行列か損失行列かを設定し、「解（候補）」のボタンをクリックする。解はまず、純粋戦略の場合から調べて行く。

純粋戦略の場合、相手が戦略を変更しなければ、自分が戦略を変更すると利得が下がるか、同じ利得となる場合が解となる。これは有限な戦略の場合、すべての戦略の組を調べればよいのでプログラムとしては簡単である。これをまず「純粋戦略」の解として表示する。

混合戦略の場合は、2 人それぞれの行列に対して、行と列を取り除きながら、補遺 2 で与えた方法で解いた答えが、取り除いた行と列の確率を 0 として与えられる。但し、利得行列か損失行列かで表示される解が異なってくる。

「解（候補）」ボタンをクリックすると、純粋戦略の解と微分を使った方法の解が表示される。前の節でも述べたように、ここで表示される解はすべての解を尽くしているわけではないことに注意してもらいたい。前節図 1 の行列を使って例を見てみよう。

例 1 プレイヤー 1：利得行列 a, プレイヤー 2：利得行列 b

行列 a, b をそれぞれ利得行列として選択し、分析実行画面の「解（候補）」と「純粋戦略解」ボタンをクリックすると図 1 のような結果を得る。

| 純粋戦略 1 | |
|---------|-------|
| プレイヤー 1 | |
| p1 | 1.000 |
| p2 | 0.000 |
| p3 | 0.000 |
| 最大化解 | 4.000 |
| プレイヤー 2 | |
| q1 | 1.000 |
| q2 | 0.000 |
| q3 | 0.000 |
| 最大化解 | 5.000 |

| 利得 \ 利得 | PL2 戦略 1 | PL2 戦略 2 | PL2 戦略 3 |
|----------|----------|----------|----------|
| PL1 戦略 1 | (4), (5) | -1, 2 | -2, 1 |
| PL1 戦略 2 | -3, (3) | 0, 2 | (3), 2 |
| PL1 戦略 3 | 3, (4) | (1), 1 | 2, 0 |

図 1 非協力非ゼロ和 2 人ゲームの解

これは純粋戦略の解だけを持つ場合の例である。「解（候補）」ボタンをクリックした際、その下のコンボボックスに解の名前がセットされ、それを選択することによって解及びプレイヤーごとに解と利得の関係を表示することができるようになる。図 2a と図 2b にそれぞれ「PL1 解（候補）」と「PL2 解（候補）」を選択した場合の結果を示す。

| Player 1 | | q1 | q2 | q3 | |
|----------|-------|--------|--------|--------|--------|
| 利得行列 | 確率 | 1.000 | 0.000 | 0.000 | |
| p1 | 1.000 | 4.000 | -1.000 | -2.000 | 4.000 |
| p2 | 0.000 | -3.000 | 0.000 | 3.000 | -3.000 |
| p3 | 0.000 | 3.000 | 1.000 | 2.000 | 3.000 |
| | | 4.000 | -1.000 | -2.000 | 4.000 |

図 2a PL1 簡易解

| Player 2 | | q1 | q2 | q3 | |
|----------|-------|-------|-------|-------|-------|
| 利得行列 | 確率 | 1.000 | 0.000 | 0.000 | |
| p1 | 1.000 | 5.000 | 2.000 | 1.000 | 5.000 |
| p2 | 0.000 | 3.000 | 2.000 | 2.000 | 3.000 |
| p3 | 0.000 | 4.000 | 1.000 | 0.000 | 4.000 |
| | | 5.000 | 2.000 | 1.000 | 5.000 |

図 2b PL2 簡易解

これを見ると相手の戦略に対する期待値（右端や下の青い部分）について、それぞれのプレイヤーが期待値の最も大きな戦略を確率 1 で選択していることが分かる。

例 2 プレイヤー 1 : 利得行列 a, プレイヤー 2 : 利得行列 g

| 戦略解 (A, G) | | 純粋戦略 1 | 純粋戦略 2 |
|------------|--|--------|--------|
| プレイヤー 1 | | | |
| p1 | | 1.000 | 0.000 |
| p2 | | 0.000 | 1.000 |
| p3 | | 0.000 | 0.000 |
| 最大化解 | | 4.000 | 3.000 |
| プレイヤー 2 | | | |
| q1 | | 1.000 | 0.000 |
| q2 | | 0.000 | 0.000 |
| q3 | | 0.000 | 1.000 |
| 最大化解 | | 5.000 | 3.000 |

| 純粋戦略解 (A, G) | | PL2 戦略 1 | PL2 戦略 2 | PL2 戦略 3 |
|--------------|--|----------|----------|----------|
| 利得 \ 利得 | | | | |
| PL1 戦略 1 | | (4), (5) | -1, 2 | -2, -2 |
| PL1 戦略 2 | | -3, 2 | 0, 1 | (3), (3) |
| PL1 戦略 3 | | 3, -1 | (1), 3 | 2, (6) |

図 3 純粋戦略解が 2 つある場合

これは異なる純粋戦略の解が 2 つある場合である。どちらも解であるが、明らかに純粋戦略 1 の解が純粋戦略 2 の解に比べて優れている。

例 3 プレイヤー 1 : 利得行列 g, プレイヤー 2 : 利得行列 h

| 戦略解 (G, H) | | 混合戦略 1 | 混合戦略 2 | 混合戦略 3 |
|------------|--|--------|--------|--------|
| プレイヤー 1 | | | | |
| p1 | | 0.286 | 0.000 | 0.500 |
| p2 | | 0.429 | 0.600 | 0.000 |
| p3 | | 0.286 | 0.400 | 0.500 |
| 最大化解 | | 2.107 | 2.500 | 2.429 |
| プレイヤー 2 | | | | |
| q1 | | 0.464 | 0.500 | 0.143 |
| q2 | | 0.214 | 0.000 | 0.857 |
| q3 | | 0.321 | 0.500 | 0.000 |
| 最大化解 | | 1.857 | 2.200 | 2.500 |

| 純粋戦略解 (G, H) | | PL2 戦略 1 | PL2 戦略 2 | PL2 戦略 3 |
|--------------|--|----------|----------|----------|
| 利得 \ 利得 | | | | |
| PL1 戦略 1 | | (5), 1 | 2, (3) | -2, 1 |
| PL1 戦略 2 | | 2, 1 | 1, 1 | 3, (3) |
| PL1 戦略 3 | | -1, (4) | (3), 2 | (6), 1 |

図 4 混合戦略の解だけがある場合

図 5a と図 5b にプレイヤー 1 とプレイヤー 2 の混合戦略の 1 番目の簡易解を示す。

| Player 1 | | q1 | q2 | q3 | |
|----------|-------|--------|-------|--------|-------|
| 利得行列 | 確率 | 0.464 | 0.214 | 0.321 | |
| p1 | 0.286 | 5.000 | 2.000 | -2.000 | 2.107 |
| p2 | 0.429 | 2.000 | 1.000 | 3.000 | 2.107 |
| p3 | 0.286 | -1.000 | 3.000 | 6.000 | 2.107 |
| | | 2.000 | 1.857 | 2.429 | 2.107 |

図 5a プレイヤー 1 の簡易解 1

| 解候補 (利得行列 H) | | | | | |
|--------------|-------|-------|-------|-------|-------|
| Player2 | | q1 | q2 | q3 | |
| ▶ 利得行列 | 確率 | 0.464 | 0.214 | 0.321 | |
| p1 | 0.286 | 1.000 | 3.000 | 1.000 | 1.429 |
| p2 | 0.429 | 1.000 | 1.000 | 3.000 | 1.643 |
| p3 | 0.286 | 4.000 | 2.000 | 1.000 | 2.607 |
| | | 1.857 | 1.857 | 1.857 | 1.857 |

図 5b プレイヤー 2 の簡易解 1

2 番目の簡易解は図 6a と図 6b となる。

| 解候補 (利得行列 G) | | | | | |
|--------------|-------|--------|-------|--------|-------|
| Player1 | | q1 | q2 | q3 | |
| ▶ 利得行列 | 確率 | 0.500 | 0.000 | 0.500 | |
| p1 | 0.000 | 5.000 | 2.000 | -2.000 | 1.500 |
| p2 | 0.600 | 2.000 | 1.000 | 3.000 | 2.500 |
| p3 | 0.400 | -1.000 | 3.000 | 6.000 | 2.500 |
| | | 0.800 | 1.800 | 4.200 | 2.500 |

図 6a プレイヤー 1 の簡易解 2

| 解候補 (利得行列 H) | | | | | |
|--------------|-------|-------|-------|-------|-------|
| Player2 | | q1 | q2 | q3 | |
| ▶ 利得行列 | 確率 | 0.500 | 0.000 | 0.500 | |
| p1 | 0.000 | 1.000 | 3.000 | 1.000 | 1.000 |
| p2 | 0.600 | 1.000 | 1.000 | 3.000 | 2.000 |
| p3 | 0.400 | 4.000 | 2.000 | 1.000 | 2.500 |
| | | 2.200 | 1.400 | 2.200 | 2.200 |

図 6b プレイヤー 2 の簡易解 2

3 番目の簡易解は図 7a と図 7b となる。

| 解候補 (利得行列 G) | | | | | |
|--------------|-------|--------|-------|--------|-------|
| Player1 | | q1 | q2 | q3 | |
| ▶ 利得行列 | 確率 | 0.143 | 0.857 | 0.000 | |
| p1 | 0.500 | 5.000 | 2.000 | -2.000 | 2.429 |
| p2 | 0.000 | 2.000 | 1.000 | 3.000 | 1.143 |
| p3 | 0.500 | -1.000 | 3.000 | 6.000 | 2.429 |
| | | 2.000 | 2.500 | 2.000 | 2.429 |

図 7a プレイヤー 1 の簡易解 3

| 解候補 (利得行列 H) | | | | | |
|--------------|-------|-------|-------|-------|-------|
| Player2 | | q1 | q2 | q3 | |
| ▶ 利得行列 | 確率 | 0.143 | 0.857 | 0.000 | |
| p1 | 0.500 | 1.000 | 3.000 | 1.000 | 2.714 |
| p2 | 0.000 | 1.000 | 1.000 | 3.000 | 1.000 |
| p3 | 0.500 | 4.000 | 2.000 | 1.000 | 2.286 |
| | | 2.500 | 2.500 | 1.000 | 2.500 |

図 7b プレイヤー 2 の簡易解 3

これらの結果を見ると、プレイヤー 1 とプレイヤー 2 は共に相手が戦略を変更しない限り、自分で戦略を変更する動機がないことが分かる。

例4 プレイヤー1：利得行列 g ，プレイヤー2：利得行列 a

| 戦略解 (G, A) | | | |
|------------|-------|-------|-------|
| | 純粋戦略1 | 混合戦略1 | 混合戦略2 |
| ▶ プレイヤー1 | | | |
| p1 | 1.000 | 0.000 | 0.500 |
| p2 | 0.000 | 0.143 | 0.500 |
| p3 | 0.000 | 0.857 | 0.000 |
| 最大化解 | 5.000 | 2.500 | 2.375 |
| プレイヤー2 | | | |
| q1 | 1.000 | 0.500 | 0.625 |
| q2 | 0.000 | 0.000 | 0.000 |
| q3 | 0.000 | 0.500 | 0.375 |
| 最大化解 | 4.000 | 2.143 | 0.500 |

| 純粋戦略解 (G, A) | | | |
|--------------|----------|--------|--------|
| 利得\利得 | PL2戦略1 | PL2戦略2 | PL2戦略3 |
| ▶ PL1戦略1 | (5), (4) | 2, -1 | -2, -2 |
| PL1戦略2 | 2, -3 | 1, 0 | 3, (3) |
| PL1戦略3 | -1, (3) | (3), 1 | (6), 2 |

図28 純粋戦略解と混合戦略解がある場合

この例は純粋戦略の解と混合戦略の解を含む場合である。

最後になるが、経済学や経営学におけるゲーム理論の講義ではモデル的な問題に当てはめた利得行列を考え、学生に行動を考えさせることに重点が置かれる。そのためこのプログラムは学生には不向きである。教員が例題などを作るときに利用できればよいと思う。

参考文献

[1] 鈴木光男，ゲーム理論入門，共立出版，1983。

補遺1 2×2 行列の均衡解のグラフを使った求め方

分析実行画面の「 2×2 解グラフ」で表示される図を見ると、すべての均衡解が求まっている。よく知られたことであるが、このプログラムで使った描き方を示しておく。

利得（損失）行列 $(\mathbf{A})_{ij} = a_{ij}$ を使った問題で、プレイヤー1が戦略1を確率 p 、プレイヤー2が戦略1を確率 q で取った場合のゲームの値 z は、以下で与えられる。

$$\begin{aligned}
 z &= a_{11}pq + a_{12}p(1-q) + a_{21}(1-p)q + a_{22}(1-p)(1-q) \\
 &= (a_{11} - a_{12} - a_{21} + a_{22})pq + (a_{12} - a_{22})p + (a_{21} - a_{22})q + a_{22} \\
 &= (c_{11}q + c_{12})p + (c_{21}q + c_{22})
 \end{aligned}$$

プレイヤー1の行列 \mathbf{A} が利得行列の場合、 z を最大化することになり、解は右側に表した範囲となる。但し、 $0 \leq q_0 \leq 1$ を超える場合はこの範囲までとする。

$$\begin{aligned}
 c_{11}q + c_{12} > 0 \text{ のとき} &\quad \rightarrow \quad p = 1 \\
 c_{11} > 0 \quad q > -c_{12}/c_{11} = q_0 &\quad (1, 1) \sim (1, q_0) \\
 c_{11} < 0 \quad q < -c_{12}/c_{11} = q_0 &\quad (1, 0) \sim (1, q_0) \\
 c_{11} = 0 \quad c_{12} > 0 &\quad (1, 0) \sim (1, 1) \\
 c_{11}q + c_{12} < 0 \text{ のとき} &\quad \rightarrow \quad p = 0 \\
 c_{11} > 0 \quad q < -c_{12}/c_{11} = q_0 &\quad (0, 0) \sim (0, q_0) \\
 c_{11} < 0 \quad q > -c_{12}/c_{11} = q_0 &\quad (0, 1) \sim (0, q_0) \\
 c_{11} = 0 \quad c_{12} < 0 &\quad (0, 0) \sim (0, 1)
 \end{aligned}$$

$c_{11}q + c_{12} = 0$ のとき $\rightarrow p = 0 \sim 1$

$$c_{11} > 0 \quad q = -c_{12}/c_{11} = q_0 \quad (0, q_0) \sim (1, q_0)$$

$$c_{11} < 0 \quad q = -c_{12}/c_{11} = q_0 \quad (0, q_0) \sim (1, q_0)$$

$$c_{11} = 0 \quad c_{12} = 0 \quad (0, 0) \sim (1, 1) \quad \text{全領域}$$

この解を組み替えると、以下のようになる。

$c_{11} > 0$ のとき

$$q < -c_{12}/c_{11} = q_0 \quad (0, 0) \sim (0, q_0)$$

$$q = -c_{12}/c_{11} = q_0 \quad (0, q_0) \sim (1, q_0)$$

$$q > -c_{12}/c_{11} = q_0 \quad (1, 1) \sim (1, q_0)$$

$c_{11} < 0$ のとき

$$q < -c_{12}/c_{11} = q_0 \quad (1, 0) \sim (1, q_0)$$

$$q = -c_{12}/c_{11} = q_0 \quad (0, q_0) \sim (1, q_0)$$

$$q > -c_{12}/c_{11} = q_0 \quad (0, 1) \sim (0, q_0)$$

$c_{11} = 0$ のとき

$$c_{12} < 0 \quad (0, 0) \sim (0, 1)$$

$$c_{12} = 0 \quad (0, 0) \sim (1, 1) \quad \text{全領域}$$

$$c_{12} > 0 \quad (1, 0) \sim (1, 1)$$

またプレイヤー 1 の行列が損失行列の場合、 $p = 1 \Leftrightarrow p = 0$ と入れ替えればよく、利得行列の場合を $r = 0$ 、損失行列の場合を $r = 1$ とすると、まとめて以下のように書ける。

$c_{11} > 0$ のとき

$$q < -c_{12}/c_{11} = q_0 \quad (r, 0) \sim (r, q_0)$$

$$q = -c_{12}/c_{11} = q_0 \quad (0, q_0) \sim (1, q_0)$$

$$q > -c_{12}/c_{11} = q_0 \quad (1-r, 1) \sim (1-r, q_0)$$

$c_{11} < 0$ のとき

$$q < -c_{12}/c_{11} = q_0 \quad (1-r, 0) \sim (1-r, q_0)$$

$$q = -c_{12}/c_{11} = q_0 \quad (0, q_0) \sim (1, q_0)$$

$$q > -c_{12}/c_{11} = q_0 \quad (r, 1) \sim (r, q_0)$$

$c_{11} = 0$ のとき

$$c_{12} < 0 \quad (r, 0) \sim (r, 1)$$

$$c_{12} = 0 \quad (0, 0) \sim (1, 1) \quad \text{全領域}$$

$$c_{12} > 0 \quad (1-r, 0) \sim (1-r, 1)$$

また、プレイヤー 2 の場合も、行列 $(\mathbf{B})_{ij} = b_{ij}$ を使って、同様にまとめられるので詳細は省略する。

補遺 2 微分を使って求める均衡解

プレイヤー 1 とプレイヤー 2 はそれぞれ m 個と n 個の戦略を持っているとする。それぞれの戦略の実行確率を p_i ($i=1, \dots, m$) と q_j ($j=1, \dots, n$) とする。プレイヤー 1 の利得（損失）行列を $(\mathbf{A})_{ij} = a_{ij}$ 、プレイヤー 2 の利得（損失）行列を $(\mathbf{B})_{ij} = b_{ij}$ とする。均衡解ではこのうちで使われない戦略も存在する。一般にプレイヤー 1 の実行確率から $m-r$ 個、プレイヤー 2 の実行確率から $n-r$ 個の戦略は使われず、実際に使われるのは双方 r 個である。しかし、これもうまく選ばないと正しい均衡解 ($0 \leq p_i, q_j < 1$) は得られない。利用しない戦略の行と列を除いて、改めてプレイヤー 1 の利得（損失）行列を $(\mathbf{A})_{ij} = a_{ij}$ 、プレイヤー 2 の利得（損失）行列を $(\mathbf{B})_{ij} = b_{ij}$ とおく。

プレイヤー 1 について、ゲームの値は以下で与えられる。

$$R = \sum_{k=1}^r \sum_{l=1}^r a_{kl} p_k q_l = \sum_{k=1}^{r-1} \sum_{l=1}^{r-1} a_{kl} p_k q_l + \sum_{k=1}^{r-1} a_{kr} p_k q_r + \sum_{l=1}^{r-1} a_{rl} p_r q_l + a_{rr} p_r q_r$$

ここに、

$$p_r = 1 - \sum_{k=1}^{r-1} p_k, \quad q_r = 1 - \sum_{l=1}^{r-1} q_l$$

これを p_i で微分する。

$$\begin{aligned} \frac{\partial}{\partial p_i} R &= \sum_{l=1}^{r-1} a_{il} q_l + a_{ir} q_r - \sum_{l=1}^{r-1} a_{rl} q_l - a_{rr} q_r \\ &= \sum_{l=1}^{r-1} a_{il} q_l + a_{ir} \left(1 - \sum_{l=1}^{r-1} q_l \right) - \sum_{l=1}^{r-1} a_{rl} q_l - a_{rr} \left(1 - \sum_{l=1}^{r-1} q_l \right) \\ &= \sum_{l=1}^{r-1} (a_{il} - a_{ir} - a_{rl} + a_{rr}) q_l + a_{ir} - a_{rr} = 0 \end{aligned}$$

これより、以下の連立方程式が得られる。

$$\begin{aligned} (\tilde{\mathbf{A}})_{ij} &= a_{ij} - a_{in} - a_{mj} + a_{mn}, \quad (\mathbf{q})_j = q_j, \quad (\mathbf{c})_i = a_{mn} - a_{in} \quad \text{として、} \\ \tilde{\mathbf{A}}\mathbf{q} &= \mathbf{c} \end{aligned}$$

プレイヤー 2 について、ゲームの値は以下で与えられる。

$$S = \sum_{k=1}^r \sum_{l=1}^r b_{kl} p_k q_l = \sum_{k=1}^{r-1} \sum_{l=1}^{r-1} b_{kl} p_k q_l + \sum_{k=1}^{r-1} b_{kr} p_k q_r + \sum_{l=1}^{r-1} b_{rl} p_r q_l + b_{rr} p_r q_r$$

これを q_j で微分する。

$$\begin{aligned} \frac{\partial}{\partial q_j} S &= \sum_{k=1}^{r-1} b_{kj} p_k - \sum_{k=1}^{r-1} b_{kr} p_k + b_{rj} p_r - b_{rr} p_r \\ &= \sum_{k=1}^{r-1} b_{kj} p_k - \sum_{k=1}^{r-1} b_{kr} p_k + b_{rj} \left(1 - \sum_{k=1}^{r-1} p_k \right) - b_{rr} \left(1 - \sum_{k=1}^{r-1} p_k \right) \\ &= \sum_{k=1}^{r-1} (b_{kj} - b_{kr} - b_{rj} + b_{rr}) p_k + b_{rj} - b_{rr} \end{aligned}$$

これより、以下の連立方程式が得られる。

$$\begin{aligned} (\tilde{\mathbf{B}})_{ij} &= b_{ij} - b_{ir} - b_{rj} + b_{rr}, \quad (\mathbf{p})_i = p_i, \quad (\mathbf{d})_j = b_{mi} - b_{mj} \quad \text{として、} \\ {}^t\mathbf{p}\tilde{\mathbf{B}} &= {}^t\mathbf{d} \quad ({}^t\tilde{\mathbf{B}}\mathbf{p} = \mathbf{d}) \end{aligned}$$

基本的に以上を解くことによって、解を求めることができる。但し、使わない戦略の確率は 0 である。