情報科学総合演習（浅野）

イメージサイエンスと数理科学

イメージサイエンスとは

イメージサイエンスは,「画像を取り扱う科学」という意味ですが，とくにデジタル画像を処理する数学・アルゴリズムを研究する分野をさします．デジタル画像は，画素（ピクセル）とよばれる点の集まりからなり，各画素にはその明るさ（輝度）に対応する数値が割り当てられています．また，カラー画像の場合は，各画素に三原色の各色の輝度に対応する数値が割り当てられています．デジタル画像の処理では，これらの数値に対してなんらかの計算を行なうことで，画像の操作を行ないます．

今日の講義では，画像の操作のうち，画像圧縮についてその原理を説明します．画像圧縮とは，画像の見かけをなるべく変えないようにして，デジタル画像のデータ量を減らす方法です．今日は，JPEG方式などで知られる「直交変換による画像圧縮」について，その原理を簡単に説明します．

アラカルトと定食

生協の西２食堂では，さまざまなメニューを１つ１つ注文して，自由に組み合わせて１回の食事をとることができます．仮に食堂に２０種類のメニューがあるとして，そこから３種類のメニューを組み合わせるとすると，その組み合わせ方は${}_{30}C_3 = 4060$通りあります．

しかし，実際には4060通りのすべての組み合わせが選ばれるわけではありません．例えば，ライス・みそ汁・カキフライという組み合わせを選ぶ人は多いでしょうが，みそ汁・豚汁・卵スープという組み合わせを選ぶ人はまずいないでしょう．つまり，実質上可能な組み合わせは4060通りもなく，もっと少ないはずです．このことは，可能な食事を表現するためには，２０種類のメニューは**冗長**であることを示しています．

そこで,仮にほとんどの人が３種類の組み合わせしか選ばないとしましょう.これらの組み合わせを,「A定食」「B定食」「C定食」としましょう．そうすると，ほとんどの客の要望は，たくさんのメニューを用意しなくても，この３種類で満たすことができます．もちろん，他の組み合わせの食事をしたい人もいるでしょうが，そういう人はわずかですから，あきらめてもらうことにします．

基底と画像圧縮

いきなり何の話かと思ったかもしれませんが,実は上の話は直交変換による画像圧縮の原理をよく表しているのです．

上の例は，何種類かの「食事」を表すのに，

(1) さまざまなメニューの自由な組み合わせで表す．
(2) いくつかのメニューをはじめから組み合わせた，いくつかの「定食」から選ぶ．

の２つの表しかたを示しています．この場合の「メニュー」や「定食」のように,「食事」を表すために組み合わされる要素になるものを**基底**とよびます．上の話は,「食事」として選ばれるメニューの組み合わせのほとんどがいくつかの種類に限られている場合,「メニュー」そのものを基底とする(1)の表しかたよりも，あらかじめいくつかのメニューを組み合わせた「定食」を基底とする(2)の表しかたのほうが，少ない数の基底で表すことができる，ということを示しています．つまり，１回の食事を(1)の基底では「ライス，みそ汁，カキフライ」といちいち述べなければならないのに対して,(2)の基底では「A定食」とひとことで表せるわけで，そのために必要なデータ量も少なくなります．

画像のデータ圧縮の場合，１つの画像を各画素の輝度をすべて並べて表すのは，(1)の基底を用いた表しかたに相当します．これに対して，世の中に頻繁に現れる画像をあらかじめ「定食」として用意しておき，それらの組み合わせ（足しあわせ）で表しておけば，(2)の表しかたで，世の中のたいていの画像がいくつかの「定食」で表せることになります．(1)から(2)への基底の変換として用いられるのが**直交変換**で，この方法によるデータ圧縮が，冒頭で述べた「直交変換による画像圧縮」です．

うどんとそばの話

そんなうまい具合に，「定食」になるくらい頻繁に現れる画像があるのでしょうか．それを見つける方法を説明するために，また食べ物の例え話をすることにします．

うどん・そばにはいろいろな種類がありますが，ここでは「きつね」と「たぬき」を考えます．一般的には，「きつね」は油揚げがのっているものであり，「たぬき」は揚げ玉がのっているもので，どちらにも「うどん」「そば」があります．つまり，うどん屋さんのメニューには「きつねうどん」「きつねそば」「たぬきうどん」「たぬきそば」の４通りがあります．そこで，これらの４通りのメニューにどれも同じくらいの数の注文があるとすると，４通りのメニューへの注文数の分布を表す図は図１のようになります（これを散布図といいます）．

ところが，大阪では「きつね」は「うどん」，「たぬき」は「そば」に決まっており，どちらも油揚げがのったものを指します．つまり，「きつねうどん」「たぬきそば」の２種類しかないわけで，その２つに同じくらいの数の注文があるとすれば，散布図は図２のようになります．

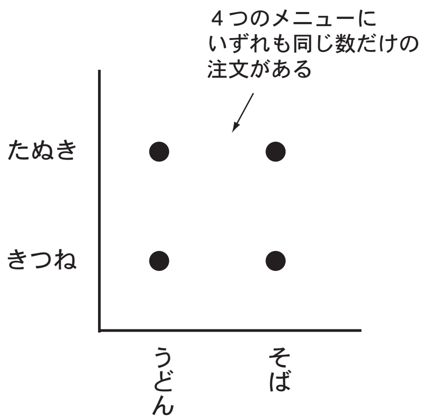

図１．一般的なうどん屋での注文の分布

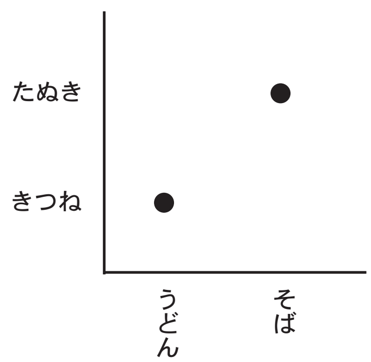

図２．大阪のうどん屋での注文の分布

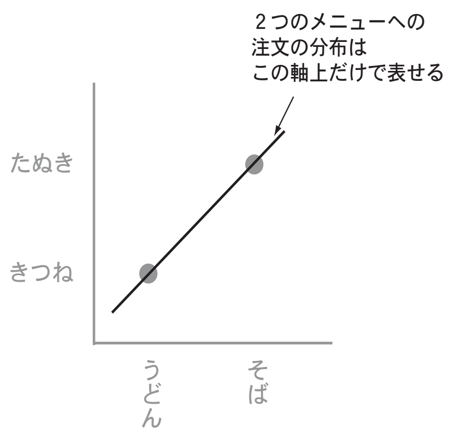

図３．２つのメニューへの注文の分布を表す軸

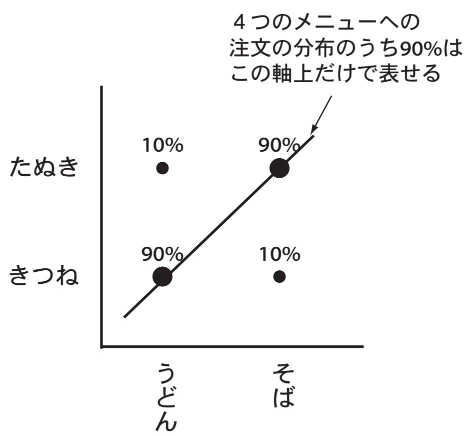

図４．４つのメニューへの注文の分布のうち
90%を表す軸

しかし，大阪での注文は２種類しかないわけですから，大阪での注文の分布を表すのに２つの軸をもつ散布図は必要ないはずです．すなわち，図３のような新しい軸を用いると，この軸上だけで注文の分布を表現することができます．

では，図４のように，「きつね」を注文する客のうち90%がうどん，10%がそばを注文し，「たぬき」を注文する客のうち90%がそば，10%がうどんを注文するとしましょう．このときは，客の注文の分布の90%は，図３と同じ軸の上で表現することができます（10%は，軸上にない注文なので，表現できません）．つまり，図３や図４の「新しい軸」が，１つの軸上で分布を最大限表現することができる，すなわち注文の大半を表すことができる「定食」に相当する基底です．

主成分分析と第１主成分

主成分分析は，前節のように，１つの軸上で分布を最大限表現することができる基底を求める方法です．

図５のような，２つの数値（変量といいます）x_1, x_2からなるデータの散布図を考えてみましょう．これらのデータについて，x_1, x_2を組み合わせて，$z_{(1)} = a_{1(1)}x_1 + a_{2(1)}x_2$という新しい変量を作ってみます*)．新しい変量$z_{(1)}$は，図中では$x_1$軸をある角度に回転させた軸上の値となります．同様に，$z_{(1)}$に直交する軸$z_{(2)}$を考えると，これも$z_{(2)} = a_{1(2)}x_1 + a_{2(2)}x_2$という新しい変量を表しています．

さて，新しい変量$z_{(1)}$が，前節で述べたような，分布を最大限表現できる変量であるとしましょう．「分布を最大限表現する」$z_{(1)}$は，$a_{1(1)}$や$a_{2(1)}$を変化させてできるさまざまな$z_{(1)}$のうち，$z_{(1)}$の分散$V(z_{(1)})$が最大であるものになります．この$z_{(1)}$を**第１主成分**といいます．

図５では，そのような$z_{(1)}, z_{(2)}$軸を描いています．散布図をこの$z_{(1)}, z_{(2)}$軸から見てみると，図から直観的に読み取れるとおり，**$z_{(1)}$と$z_{(2)}$の相関が０になっています**．

この，相関が０になるような軸を求めるため，次の**分散共分散行列**を考えます．元の変量x_1, x_2軸でみたとき，x_1の分散をs_{11}，x_2の分散をs_{22}，x_1とx_2の共分散をs_{12}とするとき，分散共分散行列は次のように定義されます．

$$\begin{pmatrix} s_{11} & s_{12} \\ s_{12} & s_{22} \end{pmatrix} \tag{1}$$

さて，データを$z_{(1)}, z_{(2)}$軸から見ると，相関が０ということは共分散が０ということですから，$z_{(1)}, z_{(2)}$軸から見たときの分散共分散行列は

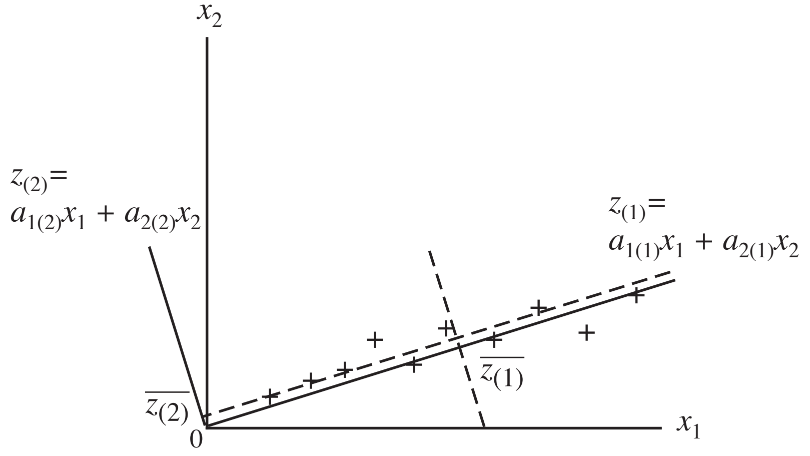

図５．主成分

*)以下，添え字の「1, 2」は元の座標軸(x)の番号，「(1), (2)」は主成分(z)の番号を表します．

$$\begin{pmatrix} \lambda_{(1)} & 0 \\ 0 & \lambda_{(2)} \end{pmatrix} \tag{2}$$

と表すことができます．このような $z_{(1)}, z_{(2)}$ 軸を求める操作は**対角化**と呼ばれています．

対角化を行うため，変量 x_1, x_2 で表されているデータを $z_{(1)}, z_{(2)}$ に変換する式を

$$z_{(1)} = a_{1(1)}x_1 + a_{2(1)}x_2,\ z_{(2)} = a_{1(2)}x_1 + a_{2(2)}x_2 \tag{3}$$

とします．このとき，a_1, a_2 と λ（それぞれ $_{(1)}$ と $_{(2)}$ の２つある）は

$$\begin{pmatrix} s_{11} & s_{12} \\ s_{12} & s_{22} \end{pmatrix}\begin{pmatrix} a_1 \\ a_2 \end{pmatrix} = \lambda\begin{pmatrix} a_1 \\ a_2 \end{pmatrix} \tag{4}$$

の解で得られることが知られています．これは**固有値問題**とよばれ，これを満たす λ は**固有値**，$\begin{pmatrix} a_1 \\ a_2 \end{pmatrix}$ は**固有ベクトル**と呼ばれています．この解は２組得られるので，２つの λ のうち大きいほうを $\lambda_{(1)}$ としそのときの固有ベクトルを $\begin{pmatrix} a_{1(1)} \\ a_{2(1)} \end{pmatrix}$，もう一方を $\lambda_{(2)}$ と $\begin{pmatrix} a_{1(2)} \\ a_{2(2)} \end{pmatrix}$ とします．

ここで，$\lambda_{(1)}, \lambda_{(2)}$ は新しい軸 $z_{(1)}, z_{(2)}$ での分散を表していますから，その大きいほうである $\lambda_{(1)}$ が第１主成分の分散になります．そして，それに対応する固有ベクトル $\begin{pmatrix} a_{1(1)} \\ a_{2(1)} \end{pmatrix}$ から(3)式で得られる $z_{(1)}$ が第１主成分となります．

画像と固有値問題

画像についても，１つの画素の輝度を１つの軸で表し，前節の方法で固有ベクトルを求めると，「定食」に相当する基底を，それぞれの画素の輝度を組み合わせた画像として求めることができるはずです．ですが，それを求めるには，分散共分散行列がわからなければなりません．分散共分散行列は，簡単に言えば「どんな画像が頻繁に現れるか」を表したものです．どんな画像が頻繁に現れるかは，何となくはわかりますが，正確に分散共分散行列で表すのは困難です．

そこで，あらかじめ基底となる画像を経験的に決めておくことにします．このような基底の中で一番よく用いられているのが，「明暗の波」，つまりいろいろな細かさの「市松模様」を基底とするものです．このような基底への変換が**フーリエ変換**で，フーリエ変換に類似の変換を用いた画像圧縮法のひとつが JPEG 方式です．実際の基底の例と圧縮の効果は，スクリーンで説明します．

今日の演習

（どちらも，少し考えてもらう必要があります）

１．文字の画像を圧縮するのに，JPEG 方式はあまり用いられないのはなぜですか．

２．プリンタの印刷見本に，「藤棚」や「満開の桜に覆われた山」などの，細かい画像が用いられるのはなぜですか．（そのプリンタが，いかに細かい部分まで再現しているのかを主張しているのではありません）

参考

今回の講義の内容については，浅野の講義「数理科学特論Ａ」（2001年度後期）シリーズ２「直交変換による画像圧縮」，「応用統計学」（2004 年度前期）第５～７回も参考にしてください．これらの講義録は，浅野の講義ウェブサイト (`http://kougi.racco.mikeneko.jp/`) の「2004 年前期以前の講義」に掲載されています．また，上記の演習の解答も，「2004 年度後期の講義」→「情報科学総合演習 II」に掲載されています．