ホームセンターの季節品の需要予測

2001MM046 松本 京子　　2001MM072 坂口 祥子
指導教員　鈴木 敦夫

1　はじめに

研究の対象としたホームセンターでは 5 万種類を超える商品を扱っている．ホームセンターでの発注は基本的に週 2 回，1 店舗でも商品数は何万種類にも及ぶため，多量の在庫が発生してしまう．なかでも季節品は売れる時期が短く，ピーク時には売上数が総数の約半分を占めることもある為予測が重要である．しかしその年の売上のピークは気候によって大きく左右される．また一定の期間を過ぎると全く売れなくなる為，最終在庫の値下げによる利益損失などの難しさも兼ね備えている．需要予測はあくまで予測であり，正確な需要を知ることは不可能である．しかし，意思決定を助ける 1 つの手段としては有効であるといえる．また，需要予測は大幅な予測から徐々にレベルアップして，状況をふまえた柔軟なモデルを作ることが必要になってくる．

2　研究方針

2.1　ホームセンターの季節品の現状

季節品の予測は大変困難である為，機会損失を回避するために，その年のピークが近づくと去年のピーク時の売上数を在庫としてもつ．そしてピークを過ぎてから徐々に発注を減らしていく．この際の問題点としては，

- 在庫費がかさむ
- 最終在庫の処分費による利益損失が起きる

利点としては

- ピークがいつ来ても品切れを起こしにくい

が挙げられる．季節品の売れ残りは値下げによる最終処分と，次の年まで商品を保管し販売する方法がある．しかし最終処分による値下げが多すぎるとそれまでの利益をマイナスしてしまい，また 1 年保管する場合は保管費がかかるという問題が挙げられる．

2.2　研究対象

夏商品 4 種類，冬商品 6 種類の売上数のデータを整理し，3 年分または 2 年分揃っているものを取り出す．その中から年ごとの価格と売上数の格差が少なく，値段が手頃で身近であり，季節変動が見やすいと思われる‘すだれ’を研究対象に用いた．

2.3　データ

データはホームセンターから頂いた週別販売数，週別売上高を用いた．店舗数は，2002 年は 87 店舗，2003 年は 99 店舗，2004 年は 113 店舗である．期間は 2002 年 4 月 1 日〜2004 年 8 月 9 日である．何年のどこの店舗のデータか分かるようにする為，3 年間のすべての店舗に順番に番号をつけたものを店舗 No. とする．

2.4　その他のデータ

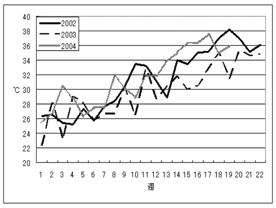

図 1　最高気温

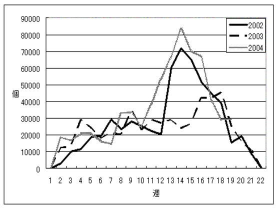

図 2　売上数

2002 年，2003 年は売上数の変化が現れる 22 週分のデータを用い，2004 年に関しては 20 週以降のデータがない為 19 週までのデータを用いる．

2002 年，2004 年はほぼ平年通りの夏でピーク時期，ピーク時期の売上数，売上総数がほぼ一致している．しかし，2003 年は冷夏であった為，はっきりとしたピークが現れず売上は伸び悩む傾向にあった．

2.5　問題へのアプローチ

本研究の目的を実際の売上数と予測との絶対誤差をできる限り少なくするモデルを考えることとする．予測方法としては次の 2 通りで行う．

- 全体の売上傾向をつかむ
- それぞれの店舗で 1 週ずつ予測する

1 つ目は最初の 2，3 週間の売上の変化を見て，ピーク時やその後のおおまかな傾向を見るのにパターン化したものを用いる．2 つ目はこれまでの週の変化に応じて次の週を予測する．

3　売上傾向の予測

まず需要予測を考える際に，その商品の売上傾向の特徴をつかむことが重要である．例えば，売れる期間，ピークの時期，ピーク時の売上数等である．これらの傾向を過去のデータを元にパターン化して考える．この章でのデータは年ごとの売上総数を用いる．

3.1　最小二乗法による近似

3.1.1　最小二乗法について

直線や曲線 (二次関数や指数関数) を仮定して現在までのデータを元に最小二乗法によって直線や曲線の係数を求める．そして求められた直線または曲線の次の点を予測値とするもの．2002 年のデータ 22 週分で予測する．

t：週の添え字 $(t = 1, 2, \cdots, 22)$
y_t：t 週目の売上数
$\hat{y}_t$：t 週目の予測値

$$\hat{y}_t = \beta_0 + \beta_1 t + \beta_2 t^2 \cdots \beta_n t^n \tag{1}$$

$$\sum_{t=1}^{22} (y_t - \hat{y}_t)^2 \longrightarrow \min \tag{2}$$

となるような $\beta_0, \beta_1, \cdots \beta_n$ を求める．

3.1.2　実行結果

2002 年のすだれの売上総数を用い，次数を 2 次，4 次，6 次と変化させ誤差の変化をみる．以下に売上数と近似した 2 次，6 次のグラフを以下に示す．

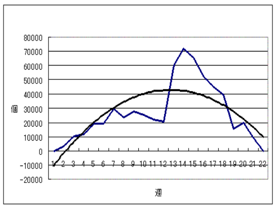

図 3　2 次式

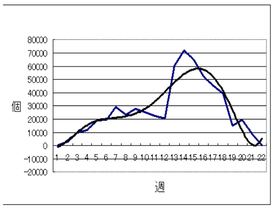

図 4　6 次式

3.1.3　考察

誤差の割合は 2 次式が 41%，4 次式が 29%，6 次式が 24% となった．次数を上げることで改善されることがわかる．しかし，売上数と予測ではピーク時期にずれが生じ，欠品を招いてしまう．また，ピーク後では売上数よりも予測が上回っているため最終在庫を増加させてしまう恐れがある．

3.2　最小二乗法に季節指数を加えた需要予測

3.2.1　最小二乗法に季節指数を加えた予測について

最小二乗法で求められた値に時系列で変化する季節指数 α_t を加え，ピーク時の急激な変化に対応できるようにする．季節指数とは，売上数の傾向が似ている年ごとに利用でき，季節的な変動をカバーできる値である．売上数の傾向が似ている 2002 年，2004 年の予測式を以下に示す．2004 年のデータを用いる為 19 週分で予測する．

t：週の添え字 $(t = 1, 2, \cdots, 19)$

i：2002 年，2004 年 (i=1，2)

y_{it}：i 年のときの t 週の売上数

$\hat{y}_{it}$：i 年のときの t 週の予測値

α_t：t 週の季節指数

$$\hat{y}_{it} = \beta_{0i} + \beta_{1i}t + \beta_{2i}t^2 + \alpha_t \tag{3}$$

$$\sum_{t=1}^{19}\sum_{i=1}^{2}(y_{it} - \hat{y}_{it})^2 \longrightarrow \min \tag{4}$$

となるような$\beta_{0i}, \beta_{1i}, \beta_{2i}, \alpha_1, \alpha_2, \cdots, \alpha_{19}$を求める．

3.2.2　実行結果

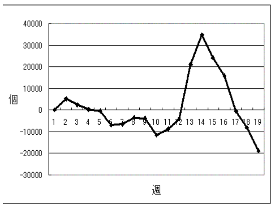

図 5　α_t の値

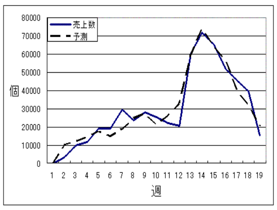

図 6　2002 年

3.2.3　考察

季節指数 α_t のグラフを見ると，ピーク時には大きな値を示していることがわかる．またピーク時の前後 5〜12 週，17〜19 週では α_t の値が 0 を下回っている．これは，最小二乗法による予測値が売上数を大きく上回っている為，その時点での過剰在庫を減らせという意味をもつ．次に, 売上数と予測のグラフを見ると，ピーク時期や売上数が的確に予測され誤差が大幅に改善されたことが分かる．α_t を加える前と後では誤差の割合が 24% から 13% となった．しかし，この結果は 2 年分のデータしか用いてない為と考えられる．

3.3　売上傾向のパターン化

最小二乗法に季節指数を加えた予測式を使いパターン化して予測を行うので，ここからは店舗ごとのデータを用いる．まず地域や傾向によりパターン化ができるかどうかを探る．店舗ごとでも立地条件，気候などにより販売力が異なるため，3 年分のデータから似ているものを取り出すことは難しい．そこで，より多くのデータを使いパターンを見つけたいので今回はすべての店舗において売上総数を 1 万と置きそれに合わせた数値を週の売上数にも掛け，売上数の増減を比較できるようにした．パターン化の方法は以下の通りである．

- ピーク時期によるもの
  各店舗のピーク時の最大値が何週目にあたるかを調べ，最大値にあたる週ごとにまとめる．次に最大値の週の上位 5 店舗，下位 5 店舗のデータを取り出し週ごとに平均を求め 1 つにまとめる．
- 県別によるもの
  気候により売上傾向が異なるため，年別県別に分け売上数の平均を求める．その際に売上傾向が大幅に異なる店舗は省く．

この 2 点の方法から求めたグラフを見比べ，売上傾向が似ているものは 1 つにまとめる．

3.3.1　実行結果

以上の手順でパターン化を行ったところ，19 種類のパターンができた．以下に 4 つのパターンを示す．

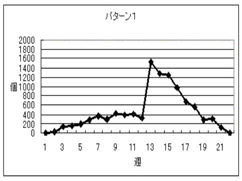

図 7　ピーク 13 上位

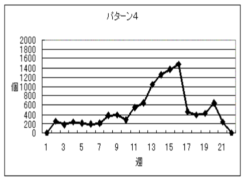

図 8　ピーク 16 上位

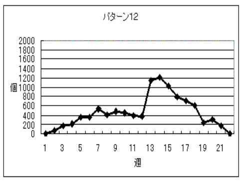

図 9　2002 愛知

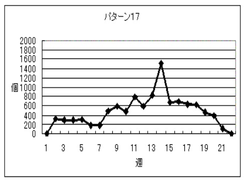

図 10　2004 愛知

3.3.2　考察

同年の愛知，岐阜，三重，静岡に似ているものが多く見られた．今回の結果では売上総数を 1 万と置いたことでパターンを見つけた為，実際の売上数と比較することは難しい．また，すべてのパターンを見つけ出したとも言い難い．しかし，パターン 12 とパターン 17 のグラフを見比べても分かるように，その年の気候や地域によって売上傾向が異なることが断定できた．5 年，10 年という長期間のデータを用いてパターン化することは有効であるといえる．

4　週ごとでの予測

需要予測でよく使われる指数平滑法での予測とパターンを用いた予測を比較，検討する．ここで扱うデータとは店舗別の週別売上数を使用する．店舗数が多いので無作為に取り出した 6 店舗について予測を行うがここでは 2 店舗の結果を載せる．なお，2004 年は 19 週までしかデータがない為予測も 19 週までとする．

4.1　指数平滑法による需要予測

4.1.1　指数平滑法について

指数平滑法とは時系列データから将来値を予測する際に利用される代表的な時系列分析手法である．移動平均を算出する加重平均法の 1 つで，在庫管理で定期発注方式における発注量予測によく用いられる．予測値 $F[t]$ は次式のようになる．

$F[t]$：t 期の予測値
$D[t]$：t 期の実際の売上数
α：平滑化定数　　$0 \leq \alpha \leq 1$

$$F[1] = 0 \tag{5}$$

$$F[t] = \alpha D[t-1] + (1-\alpha)F[t-1] \tag{6}$$

平滑化定数は，0 と 1 の間の任意の値を取る．この値が 1 に近い程今回の値を重視し，0 に近い程過去の経過を重視する．

4.1.2　実行結果

岐阜と富山の 2004 年の売上数を指数平滑法を用いて予測し，実際の売上数と比較したグラフ，累計グラフを以下に示す．今回の予測では季節品を用いるため，α の値を 0.7 と定め第 1 週目の値を 0 と置いて計算した．

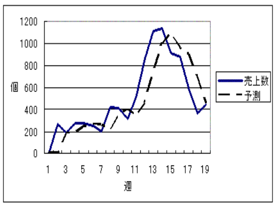

図 11　岐阜

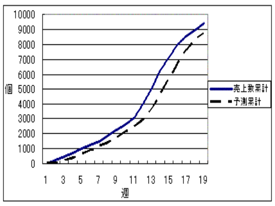

図 12　累計予測

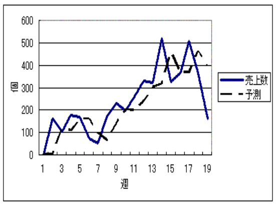

図 13　富山

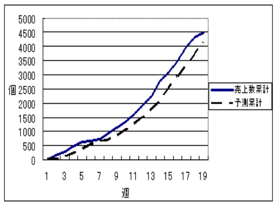

図 14　累計予測

4.1.3　考察

指数平滑法では，過去の値から未来を予測する為，予測の値が売上数を 1 週後にずらしたようなグラフになった．売上数と予測との差が大幅に増えるのは，売上数が急激に増えたり減ったりする時期であり，変化に対応しきれなかった為と考えられる．誤差の範囲は 30%～42% で 30% 台前半が多かった．

4.2　パターンを用いた需要予測

4.2.1　実際の売上数の値を返す

次に示すような定式を用いたプログラムを作って予測する．予測に使うデータとしては，2002 年，2003 年のデータを用いて 2004 年を予測することにする．入力する週の数を 3，4，5，7，10 週で試した．

【モデル化】

何週間かの実際の売上数を入力してそれと誤差が一番少なくなる店舗の予測する週の値を返す．

m：入力する週の数
n：入力を始める週
i：店舗 NO. の添え字 ($i = 1, 2, \cdots, 186$)
j：週の添え字 ($j = 1, 2, \cdots, 19$)
x_{ij}：店舗 NO.i の j 週目の売上数 (2002 年,2003 年)
y_j：予測したい店舗の j 週目の実際の売上数

これらを用いて，

$$\min_i \sum_{j=n}^{m+n} (y_j - x_{ij})^2 \tag{7}$$

を満たす $i = i^*$ を求め，$(m+n+1)$ 週目の予測値 $\hat{y}_{i^*,m+n+1}$ を求める．

【実行結果】

以下のグラフは 2004 年のデータを用い，入力する週の数を変化させた中で最も誤差が少なくなった 2 店舗の予測と実際の売上数を比較したグラフ，累積グラフである．誤差は 3 週の場合は 4～19 週，4 週の場合は 5 ～19 週を比較した．

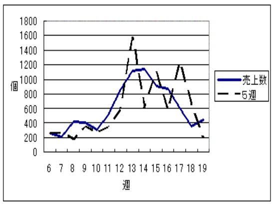

図 15　岐阜

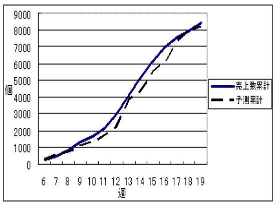

図 16　累計予測

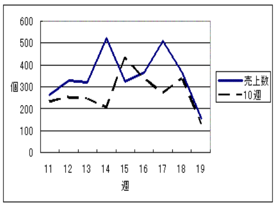

図 17　富山

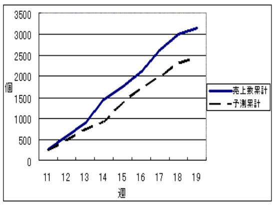

図 18　累計予測

【考察】

ピーク時期や量にずれが生じ，的確に予測することは難しかった．誤差が大幅に増える時期は指数平滑法と同様だった．誤差の範囲は 25%～41% で東海地方にしぼると 20% 台後半だった．入力する週を変化させたが今回の結果ではどの週が誤差が最小になるかという結果は得られなかった．しかし，7 週や 10 週など長くすると参照する店舗に偏りができ急激な変化がない限りその店舗を参照し続けるという結果になった．

4.2.2　売上数を比でくらべて値を返す

ここでは売上数を比で比較する方法を試した．週の和を求め，それをそれぞれの週に割り，比較する週の総和を 1 にして傾きをそろえてから比べた方が良い結果が得られるのではないかと考えた．

【モデル化】

m：入力する週の数
n：入力を始める週
i：店舗 NO. の添え字 ($i = 1, 2, \cdots, 186$)
j：週の添え字 ($j = 1, 2, \cdots, 19$)
x_{ij}：店舗 NO.i の j 週目の売上数（2002 年,2003 年）
y_j：予測したい店舗の j 週目の実際の売上数
S_i：店舗 NO.i の x_{in} から $x_{i,m+n}$ までの和
T：y_n から y_{m+n} までの和

$$X_{ij} = \frac{x_{ij}}{S_i} \qquad (j = n, \cdots, m+n) \tag{8}$$

$$Y_j = \frac{y_j}{T} \qquad (j = n, \cdots, m+n) \tag{9}$$

これらを用いて，

$$\min_i \sum_{j=n}^{m+n} (Y_j - X_{ij})^2 \tag{10}$$

を満たす $i = i^*$ を求める．

ここで比を実際の値に返す操作を行う．

比で表すと

$$\frac{x_{i^*,m+n+1}}{S_{i^*}} = \frac{\hat{y}_{i^*,m+n+1}}{T} \tag{11}$$

となるから，$\hat{y}_{i^*,m+n+1} = \dfrac{T}{S_{i^*}} x_{i^*,m+n+1}$ (12)

より $(m+n+1)$ 週目の予測値 $\hat{y}_{i^*,m+n+1}$ を求める．

【実行結果】

以下のグラフは売上数を比でくらべ，予測値を返す方法で求めた予測と実際の売上数を比較したグラフである．

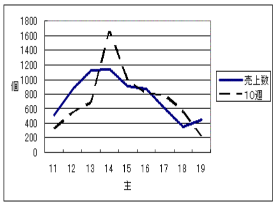

図 19　岐阜

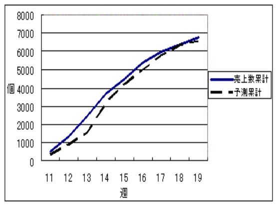

図 20　累計予測

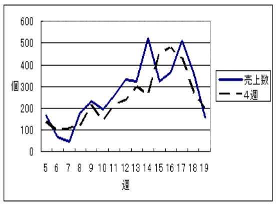

図 21　富山

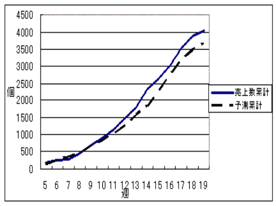

図 22　累計予測

【考察】

予測値を比で返すことにより，細かな山の変化にも少しだが対応できるようになった．誤差の範囲は 27%～55% だった．そのほかのことは前に述べたことと同様であった．比でくらべることにより誤差も少なくなり，精度がよくなると思っていたが今回の結果からはそのことを断定することができなかった．

5　おわりに

本研究では，過去のデータを元に売上数を予測することを，2 つの観点から行ってきた．データが 3 年分しかなかったため，2002 年，2003 年の 2 年分のデータをもとに，2004 年の需要予測を行った．しかし，売上数はその年の気候に左右されるところが大きいため，2 年分では当てはまりが悪く，方法があっているのかも不安な部分があった．また，売上数のグラフは似ているものを分類していくと地域別，パターン別に分けられることもわかっている．もし 5 年，10 年と長期間のデータがあればプログラムで読み込むデータを地域またはパターン別にわけてから店舗を参照すればより細かく予測できたのではないかと思う．本研究が今後の需要予測の改善方法に少しでも役に立てれば幸いである．

参考文献

[1] 久保幹雄：経営科学のニューフロンティア 8，ロジスティクス工学，朝倉書店 (2001 年).