線形回帰モデルによる保守部品の需要予測

Demand Forecasting of Service Parts Using Linear Regression

飯塚　新司　Iizuka Shinji
宗形　聡　Munakata Satoshi

保守部品の在庫管理では，部品の生産打ち切り時点でその後の保守期間分の部品をまとめて生産，あるいは調達して保有する。そのため，保守部品の生涯需要予測は製造業の在庫適正化にとって重要な技術である。(株) 日立東日本ソリューションズ（日立 TO）では，先行研究で経験分布を用いた生涯需要の予測モデルを考案した。

本報告では，経験分布モデルを拡張して，前年からの需要変化を考慮できるように一般化した線形回帰モデルを提案する。この拡張により，多くの部品で経験分布モデルよりも高精度な予測が可能になることを，実データをもとに検証する。

１．はじめに

家電や自動車などの耐久消費財では，製品修理などの保守サービスが顧客満足に大きく影響する。これらの耐久消費財の場合，保守サービスの提供期間（保守期間）は製品出荷後十数年にわたる。そのため部品の生産は保守期間の途中で終わる。そこでメーカは，部品の生産打切り時にその後の保守サービスで提供する保守部品の生涯需要を予測し，まとめて生産あるいは調達する。これをまとめ生産・まとめ調達と呼ぶ。欠品による顧客満足の低下や，過剰在庫による余分な保管コストと部品廃棄コストの発生を防ぐため，保守部品の生涯需要を高い精度で予測する技術が望まれている。

本報告では，より多くの種類の保守部品の需要予測を高精度で行えるようにするため，経験分布モデルを拡張した線形回帰モデルを提案する。このモデルでは，過去の類似部品の実績から推定した線形回帰式で前年からの需要変化を予測する。家電メーカの実データを利用して，線形回帰モデルが経験分布モデルよりも高精度な予測が可能であることを検証する。

２．経験分布モデル

先行研究で提案した経験分布モデル 1) では，まず需要の傾向が類似した部品をグループ化する。そして，製品出荷開始からの年数（経過年）ごとに，製品１台当たりの部品需要量（正規化部品需要量）の経験分布を構築する。その代表値から当該グループに属する部品の予測値を算出する。そのため予測する部品の実績を使わずに，類似部品の過去実績から予測できる，という特徴を持つ。

経験分布モデルは「同じグループ G の部品であれば，製品出荷開始から経過 t 年目の正規化部品需要量は，同一の確率分布 $F_G^{(t)}$ にしたがって発生する」という仮説に基づいている。すなわちあるグループに属する部品の経過 t 年目の正規化部品需要量を $X^{(t)}$ とすると，

$$X^{(t)} \sim F_G^{(t)}$$

である。また，$F_G^{(t)}$ の代表値を $\alpha_G^{(t)}$ とし，確率変数 $\varepsilon_G^{(t)} = X^{(t)} - \alpha_G^{(t)}$ を用いて，

$$X^{(t)} = \alpha_G^{(t)} + \varepsilon_G^{(t)} \quad \cdots (1)$$

と表すこともできる。経験分布モデルでは確率分布 $F_G^{(t)}$ の推定として，グループ内の部品の需要実績から構築する経験分布 $\hat{F}_G^{(t)}$ を用いる。

正規化部品需要量の予測値として，経験分布モデルでは $\hat{F}_G^{(t)}$ の代表値 $\alpha_G^{(t)}$ を用いる。代表値には平均値や中央値を使用する。

３．線形回帰モデルによる予測手法の提案

3.1　線形回帰モデルへの拡張

経験分布モデルでは式(1)のモデルにフィットするような需要の傾向を持つグループしか予測できない。そこで，このモデルの拡張を提案する。経過 t 年目の正規化部品需要量に対して，過去の需要量からの線形の影響を導入すると次式のようになる。

$$X^{(t)} = \sum_{j=1}^{k} \beta_{G,j}^{(t)} X^{(t-j)} + \alpha_G^{(t)} + \varepsilon_G^{(t)}$$

k は何年前の需要まで考慮するかを表す定数であり，こ

れをモデルの次数と呼ぶ。$k=0$ のとき経験分布モデルとなる。さらに正規化部品需要量に変換関数を導入し，

$$f\left(X^{(t)}\right)=\sum_{j=1}^{k}\beta_{G,j}^{(t)}\,f\left(X^{(t-j)}\right)+\alpha_G^{(t)}+\varepsilon_G^{(t)}\ \cdots\ (2)$$

とする。これは，正規化部品需要量に何らかの変換を行うことで，モデルへの適合度合いを高められる場合があるからである。例えば $f(x)=\log_{10}x$ とすることで，分散不均一性を低減できる。変換しない場合は $f(x)=x$ とする。このモデルは $f\left(X^{(t-1)}\right),\dots,f\left(X^{(t-k)}\right)$ を説明変数，$f\left(X^{(t)}\right)$ を目的変数とした重回帰モデルであり，重回帰係数 $\alpha_G^{(t)}$，$\beta_{G,1}^{(t)},\dots,\beta_{G,k}^{(t)}$ は類似部品の過去の実績から最小二乗法で推定する。$\varepsilon_G^{(t)}$ はある確率分布に従う平均 0 の誤差である。

なお，近年の部品生産打ち切りの早期化に対応するために，モデルの次数 k には 1 や 2 などの小さな値を用いることが望ましい。図 1 に $k=1$ の場合の例を示す。実線は経過 5 年目の対数部品需要量を予測する線形回帰式，点は類似部品の過去実績を表す。

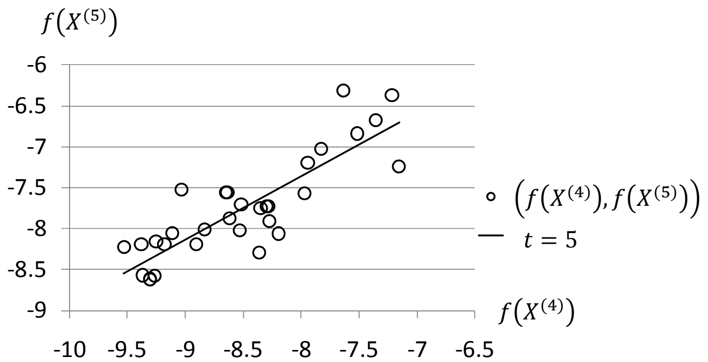

図 1 線形回帰式の例

3.2 需要量が 0 の場合に対応するための拡張

保守部品は間欠需要となることが多い。その場合，部品出荷実績が 0 のデータが多数含まれ，線形回帰式の適合度合が低くなる。そこで 3.1 節の線形回帰式では，部品出荷実績が 0 のデータは除外して重回帰係数を推定し，部品出荷台数が正の場合の部品需要量を予測する。この場合の予測値は，部品出荷台数が正の場合の条件付き期待値となる。そのため，経過 t 年目に部品出荷台数が正になる確率 $Q_G^{(t)}$ を推定して乗算することで，部品出荷台数が 0 になる場合も考慮した予測値を求める。$Q_G^{(t)}$ は

$$\hat{Q}_G^{(t)}=\frac{\text{経過}t\text{年目の部品出荷実績が正の部品数}}{\text{経過}t\text{年目の部品出荷実績が取得できる部品数}}$$

で推定する。

3.3 提案モデルの予測手順

以下，提案モデルを説明するため記号を導入する。部品の集合を P，製品の集合を M とする。部品 $p\in P$ に対して，この部品を使用する製品の集合を M_p とする。$M_p\subseteq M$ である。y 年の部品 $p\in P$ の需要量を $d_p(y)$，製品 $m\in M$ の出荷量を $s_m(y)$ とする。部品需要量の欠損を考慮するため，$d_p(y)$ が取得可能なとき $e_p(y)=1$，そうでないとき $e_p(y)=0$ とする。製品 m の出荷開始年を y_m とし，部品 p を使用する製品の出荷開始年のうち最も早い年を Y_p とする。すなわち $Y_p=\min\{y_m\mid m\in M_p\}$ とする。また，部品 p を使用するすべての製品の y 年までの累計出荷台数 $S_p(y)$ を

$$S_p(y)=\sum_{m\in M_p}\sum_{\eta=y_m}^{y}s_m(\eta)$$

で定義する。

次に，提案モデルの予測手順を述べる。

(1) 部品需要量の正規化

$e_p(y)=1$ のとき，部品 p の y 年の正規化部品需要量 $n_p(y)$ を以下の式で定義する。

$$n_p(y)=\frac{d_p(y)}{S_p(y)}$$

(2) 類似部品のグループ化

需要分布モデルの場合と同様に，需要の傾向が類似した部品を適切にグループ化する。グループ化の方法は業種，業態や扱う製品の種類によって異なるので事前の分析が必要である。例えば，次の項目によるグループ化が考えられる。

- 製品や部品の種別
- 部品の使用部位
- 部品のサイズや重量

以下，類似部品のグループは部品の集合 P の部分集合として与えられるものとする。

(3) 線形回帰式の構築

グループ G に対して，線形回帰式(2)を構築する。部品 q の経過 t 年目の正規化部品需要量を

$$x_q^{(t)}=n_q\left(Y_q+t-1\right)$$

としたとき，回帰係数 $\alpha_G^{(t)},\beta_{G,1}^{(t)},\dots,\beta_{G,k}^{(t)}$ は以下のデータセットから最小二乗法により推定する。

$$\left\{\left(f\left(x_q^{(t-j)}\right)\right)_{j=0}^{k}\,\middle|\,q\in G,\ q\neq p,\ x_q^{(t-j)}\in\mathrm{dom}(f)\ (\forall j)\right\}$$

ここで $\mathrm{dom}(f)$ は f の定義域である。例えば $f(x)=\log_{10}x$ の場合は $\mathrm{dom}(f)=\{x\mid x>0\}$ となる。

回帰係数の推定では，最低でも$(k+1)$個のデータが必要である。また，データが少ないと推定結果の信頼性が低くなる。

(4)　需要予測値の計算

予測開始の年をIとし，

$$T_p = \max\left\{t \,\middle|\, t < I - Y_p + 1,\ x_p^{(t-j)} \in \mathrm{dom}(f)\ (\forall j)\right\}$$

とする。T_pは予測開始前にk年分の部品出荷実績が取得できる最後の経過年である。ただし，部品出荷実績はfの定義域に属するとする。正規化部品需要量の予測値$\hat{x}_p^{(t)}$を，線形回帰式(2)に基づき次の漸化式で定義する。

$$\hat{x}_p^{(T_p-j)} = x_p^{(T_p-j)}\ (j = 1, \ldots, k),$$

$$\hat{x}_p^{(t)} = f^{-1}\left(\sum_{j=1}^{k} \beta_{G,j}^{(t)} f\left(\hat{x}_p^{(t-j)}\right) + \alpha_G^{(t)}\right)\ (t > T_p)$$

ここで，f^{-1}はfの逆変換である。$\hat{x}_p^{(t)}$は線形回帰式(2)の期待値であり，線形回帰モデルではこのように期待値を予測値に使用する。一方，$k = 0$の経験分布モデルでは，期待値である平均値の$\alpha_G^{(t)}$の他に中央値も使用する。

予測対象部品pのy年の部品需要量予測値$\hat{d}_p(y)$は，$t = y - Y_p + 1$として，以下の式で計算する。

$$\hat{d}_p(y) = S_p(y)\,\hat{Q}_G^{(t)}\,\hat{x}_p^{(t)}$$

４．提案手法の評価

本節では，線形回帰モデルと経験分布モデルの予測精度を比較し，線形回帰モデルへの拡張により高精度な予測が可能となることを確認する。

4.1　評価方法

実験には家電メーカA社の部品と製品の出荷データを用いた。それぞれ以下の期間のデータである。

- 部品出荷実績：2001年から2007年まで
- 製品出荷実績：1979年から2007年まで

線形回帰モデルは，次数$k = 1$，変換関数は常用対数とする。経験分布モデルは，少数の部品で観測された正規化部品需要量の異常値が予測値に影響するのを避けるため，平均値ではなく中央値を代表値に用いる。類似部品の実績がなく中央値が計算できない年があるときは，予測不可とする。部品グループは，A社のノウハウにもとづき35グループ作成して両モデルの構築に用いた。

予測精度の評価は以下の指標で行う。

(1)　絶対予測誤差率（APE）の平均

予測期間中の実績値をD，予測値をFとしたとき，APEは以下の式で定義される。

$$\mathrm{APE} = \frac{|F - D|}{D} \times 100\ (\%)$$

APE の平均は予測手法の平均的な予測性能を表し，小さいほど高精度となる。

(2)　APEが25%以下の部品件数の割合（Pred(25)）

部品の種類ごとに，全部品のうちAPEが25%以下になった部品の割合を計算し，その値で評価する。大きいほど，より多くの部品で高精度の予測が可能であるといえる。

4.2　評価結果

線形回帰モデルと経験分布モデルの予測精度の評価結果を表1に示す。集計の際，極端に大きなAPE値となった1部品により平均値が大きく変化するのを避けるため，部品の種類ごと，予測モデルごとに，該当する1部品を異常値部品として除外した。予測性能に有意差があるかどうかを確認するため，APE平均についてはt検定を行った。評価結果に合わせp値を表1に示す。

表1 予測精度の比較

部品の種類	部品の件数	APE平均（括弧内はt検定p値）		Pred(25)	
		線形回帰モデル	経験分布モデル	線形回帰モデル	経験分布モデル
室内製品外装部品	207	77.06%	69.46% (30.79%)	31.40%	10.63%
室外製品電子部品	228	44.68%	116.19% (1.52%)	46.05%	20.18%
室内製品電子部品	163	76.36%	122.80% (1.63%)	36.20%	20.25%
室内製品付属部品	41	49.90%	148.25% (0.65%)	39.02%	14.63%
総計	639	63.58%	104.79% (0.07%)	38.34%	16.74%

総計を比較すると，APE平均，Pred(25)とも経験分布モデルから線形回帰モデルへの拡張によって，より多くの部品で高い精度の予測ができているといえる。経験分布モデルでは次数が低くモデルにフィットしなかった部品でも，線形回帰モデルへの拡張によりモデルへの適合度が高まり高精度な予測ができるようになった。電子部品と付属部品では，APE平均，Pred(25)ともに経験分布モデルよりも改善している。外装部品では，予測期間中の部品出荷実績が数個程度の部品で線形回帰モデルのAPEが大きくなる傾向があり，その影響でAPE平均に有意差は見られなかった。しかし，Pred(25)が大きく向上していることから，多くの部品で予測精度が改善した

といえる。

5．おわりに

より多くの種類の保守部品の需要予測を高精度で行えるようにするため，経験分布モデルを拡張し，線形回帰モデルを考案した。家電メーカの実データを用いた実験により，モデルの拡張によって予測精度が向上することを確認した。

日立 TO では，本稿で提案した線形回帰モデルの他，製品利用条件を考慮した部品寿命分布による予測モデルも提案している [4]。モデルの前提条件や実績値への適合度合をもとに，これらのモデルを適切に選択することで，さらに多くの種類の保守部品に対応することが可能となっている。

線形回帰モデルの今後の課題として，部品グループの最適化や，グループごとの次数の最適化が挙げられる。本研究の成果を実務で適用していただきながら，顧客のノウハウやデータ分析を活用して，これらの課題を解決していく所存である。

参考文献

1) 飯塚，他：経験分布を用いた保守部品の生涯需要予測，日立 TO 技術報告，第 16 号 (2010)
2) R. L. Goodrich: “Applied Statistical Forecasting”, Business Forecast Systems, Inc. (1992).
3) J.D. Croston: “Forecasting and stock control for intermittent demands”, Operational Research Quarterly, Vol.23, No.3, pp.289-303 (1972)
4) 宗形，他：製品利用条件を考慮した部品寿命分布による保守部品需要予測と生産意思決定，日立 TO 技術報告，第 16 号 (2010)

飯塚　新司　2008 年入社
研究開発部
在庫管理，需要予測，意思決定支援技術の研究，開発
shinji.iizuka.01@hitachi-to.co.jp

宗形　聡　2003 年入社
研究開発部
統計・数理的アプローチによる業務分析，意思決定支援技術の研究開発
munakata@hitachi-to.co.jp