

成長曲線モデルを用いた大気汚染データの解析の試み

広島大・理学研究科

富田 哲治

広島大・原爆放射能医学研究所

大瀧 慈

1. はじめに

本報告は東京都の大気データに関する解析である。データは平成7年度から平成11年度までの東京都の一般大気測定局48地点と自動車排気ガス測定局35地点から得られたものである。解析目的は、各観測値の経年変化および月変化を成長曲線モデルを用いてとらえることにある。

2. 成長曲線モデル

成長曲線モデルとは、同一個体の繰り返し測定データに対して多項式などの曲線を適合し、測定間の相関を考慮した多変量モデルである。今、サンプル数を N とし、 y_j を第 j 番目の個体の繰り返し測定値の $p \times 1$ ベクトルとする。まず最初に q 個の説明変数に基づく回帰、例えば $q - 1$ 次多項式、を適合する。

$$y_i = X\beta_i + \varepsilon_j, \quad j = 1, \dots, N$$

ここで、 X は $p \times q$ の既知の個体内計画行列で $\text{rank}(X) = q$ 、 β_j は $q \times 1$ の回帰係数ベクトル、 ε_j は $p \times 1$ の個体内誤差ベクトルである。次に回帰係数ベクトル β_j に、平均の周りでの個体間の変動を考慮して、

$$\beta_j = \Xi' a_j + \nu_j, \quad j = 1, \dots, N$$

という構造を入れる。ここで、 $\Xi = [\xi_1, \dots, \xi_k]$ は $q \times k$ の未知母数行列、 a_j は $k \times 1$ の既知ベクトルで $A = [a_1, \dots, a_k]$ を個体間計画行列と呼び、 ν_j は $q \times 1$ の個体間誤差ベクトルである。また、 ε_j, ν_j は互いに独立な確率ベクトルで、 $E(\varepsilon) = 0$ 、 $\text{Cov}(\varepsilon) = \sigma^2 I_p$ 、 $E(\nu_j) = 0$ 、 $\text{Cov}(\nu_j) = \Delta \geq 0$ 、 $\text{Cov}(\varepsilon_j, \nu_j) = 0$ であるような多変量正規分布に従っているものとする。以上をまとめると、

$$y_j = X\Xi' a_j + X\nu_j + \varepsilon, \quad j = 1, \dots, N$$

となる。これは、バランスデータに対するモデルであるが、アンバランスデータに対しては Vonesh and Carter (1987) や Khatrı and Rao (1991) らにより導入されている。また、大瀧・加藤 (1996) では平均母数に特別な構造をもつ成長曲線モデルを提案している。

3. 解析結果

本報告では，主な汚染物質 SO_2 (二酸化硫黄), CO (一酸化炭素), SPM (浮遊粒子状物質), NO_2 (二酸化窒素), OX (オキシダント) について考え，解析には2次(or 3次)の多項式モデルを考えた．簡単な位置情報として，海岸線からの距離や緯度・経度を個体間の説明変数として考えた．下の表は， NO_2 のデータに関して推定されたパラメータである．第1列はパラメータの名前で $\text{East} : \text{Month}^k$ は，東経と月の k 乗の係数パラメータを意味する．

Table 1. Estimated parameters in the case NO_2

	Coe	S.E	T	P
East : Month^0	21.19	1.239	17.11	0.000
East : Month^1	2.233	0.405	5.511	0.000
North : Month^0	-13.78	2.050	-6.719	0.000
North : Month^1	-6.922	0.671	10.32	0.000

これより，南東方向で汚染リスクの上昇傾向が見られる．他の変数に関しても係数ベクトルの値から同様な解釈ができるが，十分とはいえない．特に SPM (浮遊粒子状物質) に関しては，多様な要因が複雑に影響していると考えられ，多項式では無理があると考えられる．今後の課題として，地域相関や周期性をいかにしてモデルに取り入れるかなどが考えられる．

参考文献

- [1] Khatri, C. G. and Rao, C. R. (1991). Multivariate linear models with latent variables : problems of estimation. *J. Combin. Inform. System Sci.*, **16**, 137-154.
- [2] 大瀧 慈・加藤 浩. (1996). GROWTH2 : 母数に線形構造を持つ成長曲線モデルによる統計解析のためのプログラム . 広島大学原爆放射能医学研究所年報, **37**, 181-190.
- [3] Vonesh, E. F. and Carter, R. L. (1987). Efficient inference for random coefficient growth curve models with unbalanced data. *Biometrics*, **43**, 617-628.