I-209　重回帰モデルによる安全管理 ―差の検定による解法―

熊本大学工学部土木工学教室　正員　三池亮次
同　上　正員　小林一郎
日本道路公団　神谷恵三

1. 要旨　さきにアーチダムのクラウンにおけるたわみと、たわみに影響を及ぼす水圧や堤体温度等の要因との間の重回帰モデルの構造変化が、異常であるかどうかを、重回帰における区間推定の手法を用い検出する手法を提案した。[1] この方法では、安全性の判断が即刻可能である半面、安全性を示す指標（たわみ）yより、要因X（水圧、堤体温度等）の効果を差し引いた残差が、必ずしも正規分布に従うとはいえず、理論の展開が厳密でない。　本論において適用する"重回帰における差の検定"の理論は、即時性において区間推定の手法に劣るが、観測の組数が十分に多いとき、回帰推定値が中心極限定理に従って正規分布に漸近する性質を応用し、理論的により精緻な解法を提供するものである。

2. 重回帰における区間推定の手法の適用　アーチダムのたわみδは、アーチダムのクラウンにおける堤体平均温度t_i，温度こう配α_i（断面番号$i=1,2,\cdots$），水深（$h-h_0$）の関数であって、線形回帰モデル

$$\delta=\beta_0+\sum_i\beta_{ti}t_i+\sum_i\beta_{\alpha i}\alpha_i+\sum_i\beta_{hi}(h-h_0)^i+e \tag{1}$$

ここに、βは係数，eは偏差である。ダムの変形が非可逆的で経時変化が存在するとき、上式に$\log(1+\theta/1+\theta_{0i})$（ここに$\theta$はある時点よりの日数，$\theta_{0i}$はある基準となる日数で$1+\theta\leqq 1+\theta_{0i}$のとき$\log(1+\theta/1+\theta_{0i})=0$とする）を加えればよい。顕著な経時変化は通常たん水初期に現われ、その後安定の時期が続くが、やがて劣化現象や、地震による損傷等が進行するであろう。

ある期間を品質管理でいう予備データの区画として、その後のアーチダムの挙動が予備データの期間と比べて異常であるかどうかを、区間推定の手法を用い検出することについては先に発表のとおりである。すなわち、ある観測変量yと、p個の要素より成る説明変数マトリックスXの間に線形回帰モデル

$$y=X\beta+e \tag{2}$$

が成立するものとする。ここにβは回帰係数，eは偏差ベクトルである。通常の重回帰分析において、偏差eの母平均が$E[e]=0$，母分散$V[e_\alpha]=\sigma^2$, $\alpha=1,2,\cdots n$（これを分散共分散マトリックス$Cov[e]=\sigma^2 I$と表わす）で偏差eの要素e_αは互に独立に正規分布に従うと仮定する。

予備データの期間において式(2)を設定し、回帰係数の推定値$\hat{\beta}=S^{-1}X^Ty$，残差平方和$S_E=y^Ty-\hat{\beta}^TX^Ty$が得られたものとする。次に安全性を検定しようとする区画に対しての説明変数を$X_0^T=[x_0^{(1)}\ x_0^{(2)}\cdots\ x_0^{(n)}]$，観測変量を$y_0$とし、回帰係数は式(2)と同じで

$$y_0=X_0\beta+e_0 \tag{3}$$

とする。予備期間の偏差eと安全性を検定しようとする管理期間の偏差e_0とは、独立であるという仮定の下で、残差$v_0=y_0-X_0\hat{\beta}$の分散共分散マトリックス$Cov[v_0]=(I+X_0S^{-1}X_0^T)\sigma^2$である。$e$および$e_0$が正規分布に従い、$e\sim n(0,\sigma^2S^{-1})$であるとき、管理期間のある時点における信頼係数95%の信頼区間は、

$$-t(n-p;0.05)\leqq\frac{v_{0\alpha}}{\sqrt{(1+D_0^{\alpha\alpha})\frac{S_E}{n-p}}}\leqq t(n-p;0.05) \tag{4}$$

ここに $D_0^{\alpha\alpha}=x_{0\alpha}^T\mathcal{S}^{-1}x_{0\alpha}$ で $t(n-p;0.05)$ は、両側5%の t 値であり、これは、標準化された残差 $v_{0\alpha}/\sqrt{(1+D_0^{\alpha\alpha})S_E/n-p}$ の管理限界を与えることになる。すなわち式(4)は重回帰における区間推定による安全管理の基礎式で、その適用例についても既に発表のとおりである。式(4)は偏差 e, e_0 が正規分布に従わないとき、厳密には成立しない。

4. 重回帰における差の検定理論の適用　管理区間の回帰係数が、予備期間の回帰係数 β とは異なっており、β_0 であるという仮設を設ける。その線形回帰モデルは

$$y_0=X_0\beta_0+e_0 \tag{5}$$

であり、

$$\hat{\beta}_0=\mathcal{S}_0^{-1}X_0^Ty_0=\beta_0+\mathcal{S}_0^{-1}X_0^Te_0 \quad , \quad \mathcal{S}_0=X_0^TX_0 \tag{6}$$

で、この $\hat{\beta}_0$ を用いた回帰関数 $\mu_0=X_0\beta_0$ の推定値は $\hat{\mu}_0=X_0\hat{\beta}_0$ である。

式(5)の e_0 が、必ずしも正規分布に従わなくても、観測の組数 $n_0\to\infty$ において、回帰係数の推定値 $\hat{\beta}_0$ は重回帰における中心極限定理に従って正規分布に漸近する。すなわ $n_0\to\infty$ において $\hat{\beta}_0-\beta_0\sim n(0,\sigma^2\mathcal{S}_0^{-1})$ であることが証明される。偏差 e についても同様である。

もし、β_0 は、実は予備データの回帰係数 β と差がないと仮定すると、管理区間における回帰関数の推定値は $\hat{\mu}_0'=X_0\hat{\beta}$ であってもよい。$\hat{\beta}_0$ と $\hat{\beta}$ の差が顕著のとき、$\hat{\mu}_0$ と $\hat{\mu}_0'$ の差が顕著となるであろう。

さて

$$\left.\begin{aligned}&E[\hat{\mu}_0]=X_0E[\hat{\beta}_0]=\mu_0\\&Cov[\hat{\mu}_0]=E[(\hat{\mu}_0-\mu_0)(\hat{\mu}_0-\mu_0)^T]=X_0\mathcal{S}_0^{-1}X_0^T\sigma_0^2 \;,\quad \because\; Cov[\hat{\beta}_0]=\mathcal{S}_0^{-1}X_0^TCov[e_0]X_0\mathcal{S}_0^{-1}=\mathcal{S}_0^{-1}\sigma_0^2\end{aligned}\right\} \tag{7}$$

偏差 e_0 が正規分布に従わなくても、中心極限定理によって $\hat{\beta}_0$ したがって正規分布の再生性によって $\hat{\mu}_0=X_0\hat{\beta}_0$ は正規分布に従う。これより

$$\hat{\mu}_0\sim n(\mu_0,X_0\mathcal{S}_0^{-1}X_0^T\sigma_0^2)\;,\quad X_0\mathcal{S}_0^{-1}X_0^T\text{の対角要素を }D^{\alpha\alpha}\text{とする。} \tag{8}$$

同様に

$$\hat{\mu}_0'\sim n(\mu_0',X_0\mathcal{S}^{-1}X_0^T\sigma^2)\;,\quad X_0\mathcal{S}^{-1}X_0^T\text{の対角要素を }D'^{\alpha\alpha}\text{とする。} \tag{9}$$

したがって、正規分布の再生性に従って

$$N=\frac{(\hat{\mu}_0'-\mu_0')-(\hat{\mu}_0-\mu_0)}{\sqrt{(D_0'^{\alpha\alpha}\sigma_0^2+D_0^{\alpha\alpha}\sigma^2)}}\sim n(0,1^2) \tag{10}$$

また、$S_E/\sigma^2\sim\chi^2(n-p)$, $S_{E0}/\sigma_0^2\sim\chi^2(n_0-p)$ である。したがって $S_E/\sigma^2+S_{E0}/\sigma_0^2\sim\chi^2(n+n_0-2p)$

予備データの期間と管理期間における回帰係数 $\hat{\beta}_0$ と $\hat{\beta}$ の間に差がなく $\beta_0=\beta$ の帰無仮設を設けると、$\mu_0'=\mu_0$, $u_0'=u_0$ また $\sigma_0^2=\sigma^2$ であり、

$$T=\frac{\hat{\mu}_0'-\hat{\mu}_0}{\sqrt{(D_0'^{\alpha\alpha}+D_0^{\alpha\alpha})V}}\sim t(n+n_0-2p)\;,\quad V=\frac{S_E+S_{E0}}{n+n_0-2p} \tag{11}$$

したがって、信頼係数が95%の T の信頼区間は

$$|T|\leqq t(n+n_0-2p;0.05) \tag{12}$$

上式の右辺は、左辺の T の管理限界を与えることになる。

5. 適用例　Aアーチダムや Hアーチダムのたわみによる安全管理に適用した。3. および 4. による方法を併用することが好ましい。また、ダム以外にも橋梁や、環境指標による環境保全の問題にも応用することができる。　参考文献 1) 三池他：“重回帰モデルによるダムの安全管理”土木学会第36回年次学術講演会・昭和56年10月