報　文

水位―流量関係から推定される流量の精度について（その1）

Accuracy of Discharge Estimates from Water Level-Discharge Relationships (Part 1)

中津川　誠*　　星　　清**

Makoto NAKATSUGAWA and Kiyoshi HOSHI

水位や流量などの水文観測データは防災，水資源計画，さらには水辺環境の保全のためなどに使われる重要な資料である。

水位は時間的に連続した観測が可能であるが，流量についてはそれが不可能であり，その時系列的な値を知るには水位～流量関係を用いての推定を行うことになる。

本報告では，従来用いられている水位と流量の平方根値に直線回帰式をあてはめたものと，今回新たに提案する水位と流量に2次曲線回帰式をあてはめたものから推定された流量値の精度を統計的評価法によって比較した。これによって，いわゆる実測流量に含まれる誤差（精度）を定量的に表わすとともに，後者の方が特にデータに不足を生ずる高水時の流量について精度の高い推定を行い得ることを示した。また，高水時の流量を精度よく外挿するために必要となる観測流量の規模について示唆を与えた。

≪水位―流量関係；流量推定精度；最小二乗法；標準誤差≫

Discharge is an important paramater in the planning and management of water resources. In general, discharge data are calibrated from water levels with data of the water level (H) and discharge (Q), but little is known about the accuracy of discharge values estimated in this manner.

In this paper, a statistical evaluation was made between two methods of evaluating the water level-discharge ($H-Q$) relationship. The first method uses a conventional formula of $\sqrt{Q}=a_0+a_1H$, the second the ralation $Q=a_0+a_1H+a_2H$ where the parameters of a_0, a_1 and a_2 are estimated by a least-squares method. The accuracy of discharge estimates with the two methods are quantitatively evaluated with statistical methods. The second approach results in smaller error variances than the first, particularly where measurements of flood flows are few. Numerical experiments were conducted to determine the number of discharge measurements in high flow conditions necessary to attain the confidence limits of discharge estimates. Some examples of applications using actual measurement data in three river watersheds are presented.

Keywords: water level-discharge relationship, accuracy of discharge estimate, least-squares method, sampling property.

*環境研究室員　**前河川研究室長　現石狩川開発建設部札幌河川事務所長

まえがき

水位や流量などの水文観測データは防災，水資源計画などの重要な資料である。水位は洪水時における予警報や水防活動の情報として，渇水時における取水量や取水位の情報として用いられる。また，流量は治水計画や水資源計画策定のため，さらに近年では，生態系や水辺の環境保全上の情報としても重要度を増している。

この中で，水位に関しては時間的に連続した観測が可能であるが，流量に関しての直接的な観測はそれが不可能であり，その時系列的な値を知るには水位～流量($H-Q$)関係から換算することになる。この $H-Q$ 関係については，慣例的に水位と流量の平方根，すなわち H と$\sqrt{Q}$の間に直線関係がなりたつとして求められている。厳密には流量に係わる要因としては，水位のほか河道断面積，エネルギー勾配や河床の粗度などが考慮されねばならないが，これらの影響は水位のみとの関係の中に丸められ，さらに，データ処理上の簡便さから $H-\sqrt{Q}$の直線関係が用いられているというのが実情である。したがって，河道断面が複雑な形をしていたり，水位によって流水断面積が著しく変化するような場合，あるいは河床変動の激しい個所では当然一律の $H-Q$ 関係をつくりがたい。また，感潮区間や洪水時のように水位～流量が二価的関係をもつ場合もしかりである。このような場合にはある期間，ある水位の範囲で $H-Q$ 関係を使い分けて対処している。しかしながら，このような不連続的な運用をするほど，個々の関係を構築するために適用される実測データ自体も分断されることから，構成データ数の減少に伴う推定精度の低下を招くことも予想される。これは，特にデータが不足しがちとなる高水時の場合に懸念されるであろう。

このような背景をもとに，本報告では既往の $H-\sqrt{Q}$関係に直線回帰式をあてはめた場合と，その改良案として $H-Q$ 関係に直接２次曲線回帰式をあてはめ，低水から高水まで全データを使用して適用範囲を拡大した場合の流量の推定精度を比較検討するものである。

また，高水時の流量において，$H-Q$ 関係の外挿からどの程度の精度をもつ流量推定が可能かを既往データの間引きと精度評価手法をもって言及していく。

1.　H－Q 関係の推定方法

現在，$H-Q$ 関係の求め方は $H-\sqrt{Q}$関係が直線関係にあるとして，$H-\sqrt{Q}$の関係を最小二乗法で求めている。回帰式を次式のように求めるとする。

$$\sqrt{Q_i}=a_0+a_1H_i \qquad \text{— (1.1)}$$

ここで，Q_i；i 番目に属する流量値，H_i；i 番目に属する水位値。

上式の左辺と右辺の差の平方和は，

$$S=\sum_{i=1}^{n}\left(\sqrt{Q_i}-a_0-a_1H_i\right)^2 \qquad \text{— (1.2)}$$

となり，最小二乗法の定義より，(1.2)式の S を最小とする a_0，a_1を求めればよい。

$$\frac{\partial S}{\partial a_0}=-2\left(\sum_{i=1}^{n}\sqrt{Q_i}-a_0n-a_1\sum_{i=1}^{n}H_i\right)=0$$

$$\frac{\partial S}{\partial a_1}=-2\left(\sum_{i=1}^{n}H_i\sqrt{Q_i}-a_0\sum_{i=1}^{n}H_i-a_1\sum_{i=1}^{n}H_i^2\right)=0 \qquad \text{— (1.3)}$$

上式を解くための式形をマトリックス表示すると，

$$\begin{pmatrix} n & \sum_{i=1}^{n}H_i \\ \sum_{i=1}^{n}H_i & \sum_{i=1}^{n}H_i^2 \end{pmatrix}\begin{pmatrix} a_0 \\ a_1 \end{pmatrix}=\begin{pmatrix} \sum_{i=1}^{n}\sqrt{Q_i} \\ \sum_{i=1}^{n}H_i\sqrt{Q_i} \end{pmatrix} \qquad \text{— (1.4)}$$

となり，これを a_0，a_1について解いてやれば回帰式が求まるわけである。(1.4) 式のような式を正規方程式というが，この正規方程式の一般形については補遺～１を参照されたい。(1.4) 式を a_0，a_1について解いた結果を次に示す。

$$a_0=\frac{\sum_{i=1}^{n}\sqrt{Q_i}\sum_{i=1}^{n}H_i^2-\sum_{i=1}^{n}H_i\sum_{i=1}^{n}H_i\sqrt{Q_i}}{n\sum_{i=1}^{n}H_i^2-\left(\sum_{i=1}^{n}H_i\right)^2}$$

$$a_1=\frac{n\sum_{i=1}^{n}H_i\sqrt{Q_i}-\sum_{i=1}^{n}H_i\sum_{i=1}^{n}\sqrt{Q_i}}{n\sum_{i=1}^{n}H_i^2-\left(\sum_{i=1}^{n}H_i\right)^2} \qquad \text{— (1.5)}$$

この際の計算はルーチンに従って行えばいたって簡単なものであり，そのような理由から，実用面ではこれが使われてきていた。しかしながら，まえがきでも述べたように，現実には $H-\sqrt{Q}$が直線関係にならない場合もあり，このような場合に従来の方法で $H-Q$ 関係を求めることは流量推定の精度低下につながる。そこで，$H-\sqrt{Q}$関係に無理やり直線関係をあてはめてやるのではなく，$H-Q$ 関係に直接２次曲線をあてはめることを考える。$H-Q$ 関係を表わす２次曲線も最小二乗法を用いて求めるわけだが，先述した直線に対する最小二乗法に比べて計算には若干の手間が必要とはいえ，計算機の発達した現在においてその作業は雑作もないであろう。

２次曲線回帰式の最小二乗法による算定についても，

直線回帰式と同様に正規方程式の利用を考えていく。回帰式を以下のように仮定する。なお，このとき2次曲線式を H と H^2の線形結合式とみなして，正規方程式の導入を図るわけである。

$$Q_i = a_0 + a_1 H_i + a_2 H_i^2 \qquad -(1.6)$$

2次曲線回帰式の場合の正規方程式は，以下のとおりとなる。

$$\begin{pmatrix} n & \sum_{i=1}^{n} H_i & \sum_{i=1}^{n} H_i^2 \\ \sum_{i=1}^{n} H_i & \sum_{i=1}^{n} H_i^2 & \sum_{i=1}^{n} H_i^3 \\ \sum_{i=1}^{n} H_i^2 & \sum_{i=1}^{n} H_i^3 & \sum_{i=1}^{n} H_i^4 \end{pmatrix} \begin{pmatrix} a_0 \\ a_1 \\ a_2 \end{pmatrix} = \begin{pmatrix} \sum_{i=1}^{n} Q_i \\ \sum_{i=1}^{n} H_i Q_i \\ \sum_{i=1}^{n} H_i^2 Q_i \end{pmatrix} \qquad -(1.7)$$

これを a_0，a_1，a_2について解くと，以下の結果が得られる。

$$a_0 = \overline{Q} - a_1 \overline{H} - a_2 \overline{H}^2$$
$$a_1 = \frac{S_{HQ} S_{H^2H^2} - S_{H^2Q} S_{HH^2}}{S_{HH} S_{H^2H^2} - (S_{HH^2})^2}$$
$$a_2 = \frac{S_{HH} S_{H^2Q} - S_{HH^2} S_{HQ}}{S_{HH} S_{H^2H^2} - (S_{HH^2})^2} \qquad -(1.8)$$

ここで，

$$\overline{Q} = \frac{1}{n}\sum_{i=1}^{n} Q_i\ ,\quad \overline{H} = \frac{1}{n}\sum_{i=1}^{n} H_i\ ,\quad \overline{H}^2 = \frac{1}{n}\sum_{i=1}^{n} H_i^2$$
$$S_{HH} = \sum_{i=1}^{n} (H_i - \overline{H})^2,\quad S_{HH^2} = \sum_{i=1}^{n} (H_i - \overline{H})(H_i^2 - \overline{H}^2)$$
$$S_{H^2H^2} = \sum_{i=1}^{n} (H_i^2 - \overline{H}^2)^2,\quad S_{HQ} = \sum_{i=1}^{n} (H_i - \overline{H})(Q_i - \overline{Q})$$
$$S_{H^2Q} = \sum_{i=1}^{n} (H_i^2 - \overline{H}^2)(Q_i - \overline{Q})$$

このように，(1.8)式で得られた a_0，a_1，a_2から(1.6)式のような2次曲線回帰式が得られる。

この両者の効用については，水位から推定される流量の値がどれだけの信頼性をもつか，いい換えれば誤差を含むかを評価して下されるであろう。このことについては，次章以下に記す。

2. H－Q 関係の信頼性評価手法

常識的に考えれば，$H-Q$ 関係をつくるための実測データが少ないと流量を推定するための信頼性は低下するであろう。また，ある関係（直線関係や2次曲線など）をあてはめた場合，そのときにデータの相関が小さければ信頼性が低くなることも直観的にわかるであろう。このような評価を，正規分布をベースとしたデータの統計的処理によって行う方法について以下で述べる。

2.1 直線回帰式の信頼区間

$H-\sqrt{Q}$関係のように，直線回帰を仮定した場合に回帰式は以下のように表わすことができる。

$$\sqrt{Q} = \hat{a}_0 + \hat{a}_1 H \qquad -(2.1)$$

ここで，$\hat{a}_0$，$\hat{a}_1$は回帰係数の推定値で，以後 ^ は推定量を表わすものとする。

回帰式から求める$\sqrt{Q}$の分散推定値は，以下のようになる（補遺～2参照）。

$$\hat{V}(\sqrt{Q}) = \left(\frac{\partial(\sqrt{Q})}{\partial \hat{a}_0}\right)^2 \hat{V}(\hat{a}_0) + \left(\frac{\partial(\sqrt{Q})}{\partial \hat{a}_1}\right)^2 \hat{V}(\hat{a}_1) + 2\left(\frac{\partial(\sqrt{Q})}{\partial \hat{a}_0}\right)\left(\frac{\partial(\sqrt{Q})}{\partial \hat{a}_1}\right) \hat{Cov}(\hat{a}_0, \hat{a}_1)$$
$$= \hat{V}(\hat{a}_0) + H^2 \hat{V}(\hat{a}_1) + 2H \hat{Cov}(\hat{a}_0, \hat{a}_1) \qquad -(2.2)$$

ここで，$\hat{V}(x)$は x の分散の推定値，$\hat{Cov}(x, y)$ は x と y の共分散の推定値を表わすものである。

最終的には流量は$\sqrt{Q}$を二乗して Q の形で表わすので，その過程における分散を求める必要がある。すなわち，

$$Q = (\sqrt{Q})^2 \qquad -(2.3)$$

なる操作が必要である。よって，

$$\hat{V}(Q) = \left(\frac{\partial Q}{\partial(\sqrt{Q})}\right)^2 \hat{V}(\sqrt{Q}) = (2\sqrt{Q})^2 \hat{V}(\sqrt{Q})$$
$$= 4Q\hat{V}(\sqrt{Q}) \qquad -(2.4)$$

となり，これと (2.2) 式から，

$$\hat{V}(Q) = 4Q\ \{\hat{V}(\hat{a}_0) + H^2 \hat{V}(\hat{a}_1) + 2H\ \hat{Cov}(\hat{a}_0,\ \hat{a}_1)\} \qquad -(2.5)$$

が得られる。流量に関する誤差が正規分布に従うと仮定すれば，ある水位の値 H に対して95%の推定信頼区間は次式で求められる。

$$(\hat{a}_0 + \hat{a}_1 H)^2 \pm 1.96\sqrt{\hat{V}(Q)} \qquad -(2.6)$$

これによって，$H-\sqrt{Q}$直線回帰式から推定される流量の信頼区間が算出される。ところで，(2.5) 式中の回帰係数の分散，共分散を求めておく必要があるが，その算出法を以下に記す。

最初に回帰係数の推定値 $\hat{a}_0$, $\hat{a}_1$に着目すれば，(1.4)式で示される正規方程式の係数行列の逆行列に流量データの分散をかけたものが分散，共分散行列に等しいことがわかっている。すなわち，以下のようになる。

$$\begin{pmatrix} n & \sum_{i=1}^{n} H_i \\ \sum_{i=1}^{n} H_i & \sum_{i=1}^{n} H_i^2 \end{pmatrix}^{-1} \hat{\sigma}^2 = \begin{pmatrix} \hat{V}(\hat{a}_0) & \hat{Cov}(\hat{a}_0, \hat{a}_1) \\ \hat{Cov}(\hat{a}_0, \hat{a}_1) & \hat{V}(\hat{a}_1) \end{pmatrix} \quad \text{— (2.7)}$$

ここで，$\hat{\sigma}^2$は流量データの分散の不偏推定値を表わしており，この場合，

$$\hat{\sigma}^2 = \frac{1}{n-2} \sum_{i=1}^{n} (\sqrt{Q_i} - \hat{a}_0 - \hat{a}_1 H_i) \quad \text{— (2.8)}$$

となる。

なお，(2.7)式のような関係が成立することの証明は，補遺～3を参照されたい。

(2.7)式を展開すると分散，共分散は以下のように得られる。

$$\hat{V}(\hat{a}_0) = \left\{ \frac{1}{n} + \frac{\overline{H}^2}{\sum_{i=1}^{n} (H_i - \overline{H})^2} \right\} \hat{\sigma}^2 \quad \text{— (2.9)}$$

$$\hat{V}(\hat{a}_1) = \frac{1}{\sum_{i=1}^{n} (H_i - H)^2} \hat{\sigma}^2 \quad \text{— (2.10)}$$

$$\hat{Cov}(\hat{a}_0, \hat{a}_1) = -\frac{\overline{H}}{\sum_{i=1}^{n} (H_i - \overline{H})^2} \hat{\sigma}^2 \quad \text{— (2.11)}$$

ここで，$\overline{H}$ は水位データの平均値で，$\overline{H} = \frac{1}{n} \sum_{i=1}^{n} H_i$ である。

これら (2.9) ～(2.11) 式を (2.5) 式に代入すれば，流量の分散推定値が求まる。再度整理して (2.5) 式を記すと，ある水位値 H_0に対して推定される流量 Q_0の分散（推定値）は，

$$\hat{V}(Q_0) = 4(\hat{a}_0 + \hat{a}_1 H_0) \left\{ \frac{1}{n} + \frac{(H_0 - \overline{H})^2}{\sum_{i=i}^{n} (H_i - \overline{H})^2} \right\} \hat{\sigma}^2 \quad \text{— (2.12)}$$

95%の信頼区間は，

$$(\hat{a}_0 + \hat{a}_1 H_0) \pm 1.96 \sqrt{\hat{V}(Q_0)} \quad \text{— (2.13)}$$

となる。

(2.12)，(2.13) 式はデータ個数 n が少ないほど，観測データのばらつき（分散）が大きいほど，また水位，すなわち流量が大きくなるほど流量の分散推定値が大きくなり，信頼性が低下することを表わしている。

2.2　2次曲線回帰式の信頼区間

$H-Q$ 関係に2次曲線をあてはめた場合，回帰式は次のような形で表わすことができる。

$$Q = \hat{a}_0 + \hat{a}_1 H + \hat{a}_2 H^2 \quad \text{— (2.14)}$$

ここで，$\hat{a}_0$，$\hat{a}_1$，$\hat{a}_2$は回帰係数の推定値。

この2次曲線で推定される流量についても，前述したのと同様の方法論で信頼性を評価できる。最初に，(2.14)式で求まる流量 Q の分散推定値を示す（補遺～2参照）。

$$\begin{aligned} \hat{V}(Q) = & \left(\frac{\partial Q}{\partial \hat{a}_0}\right)^2 \hat{V}(\hat{a}_0) + \left(\frac{\partial Q}{\partial \hat{a}_1}\right)^2 \hat{V}(\hat{a}_1) \\ & + \left(\frac{\partial Q}{\partial \hat{a}_2}\right)^2 \hat{V}(\hat{a}_2) \\ & + 2\left(\frac{\partial Q}{\partial \hat{a}_0}\right)\left(\frac{\partial Q}{\partial \hat{a}_1}\right) \hat{Cov}(\hat{a}_0, \hat{a}_1) \\ & + 2\left(\frac{\partial Q}{\partial \hat{a}_1}\right)\left(\frac{\partial Q}{\partial \hat{a}_2}\right) \hat{Cov}(\hat{a}_1, \hat{a}_2) \\ & + 2\left(\frac{\partial Q}{\partial \hat{a}_2}\right)\left(\frac{\partial Q}{\partial \hat{a}_0}\right) \hat{Cov}(\hat{a}_2, \hat{a}_0) \\ = & \hat{V}(\hat{a}_0) + H^2 \hat{V}(\hat{a}_1) + H^4 \hat{V}(\hat{a}_2) \\ & + 2H \hat{Cov}(\hat{a}_0, \hat{a}_1) + 2H^3 \hat{Cov}(\hat{a}_1, \hat{a}_2) \\ & + 2H^2 \hat{Cov}(\hat{a}_2, \hat{a}_0) \end{aligned} \quad \text{— (2.15)}$$

前述したのと同様に，(1.6) 式の2次曲線正規方程式の係数行列の逆行列に流量データの分散をかけたものが分散，共分散行列に等しいことを利用する（補遺～3参照）。

$$\begin{pmatrix} n & \sum_{i=1}^{n} H_i & \sum_{i=1}^{n} H_i^2 \\ \sum_{i=1}^{n} H_i & \sum_{i=1}^{n} H_i^2 & \sum_{i=1}^{n} H_i^3 \\ \sum_{i=1}^{n} H_i^2 & \sum_{i=1}^{n} H_i^3 & \sum_{i=1}^{n} H_i^4 \end{pmatrix}^{-1} \hat{\sigma}^2$$

$$= \begin{pmatrix} \hat{V}(\hat{a}_0) & \hat{Cov}(\hat{a}_0, \hat{a}_1) & \hat{Cov}(\hat{a}_0, \hat{a}_2) \\ \hat{Cov}(\hat{a}_1, \hat{a}_0) & \hat{V}(\hat{a}_1) & \hat{Cov}(\hat{a}_1, \hat{a}_2) \\ \hat{Cov}(\hat{a}_2, \hat{a}_0) & \hat{Cov}(\hat{a}_2, \hat{a}_1) & \hat{V}(\hat{a}_2) \end{pmatrix} \quad \text{— (2.16)}$$

ここで $\hat{\sigma}^2$は水位・流量データの分散の不偏推定値を表わしており，この場合，

$$\hat{\sigma}^2 = \frac{1}{n-3}\sum_{i=1}^{n}(Q_i - \hat{a}_0 - \hat{a}_1 H_i - \hat{a}_2 H_i^2) \qquad -(2.17)$$

となる。

(2.16) 式を展開すると分散，共分散がそれぞれ求まるが，計算は複雑であるので結果のみを示す。この正規方程式の係数マトリックスを A とすれば，(2.16) 式を解くには A の逆行列 A^{-1}を求めればよい。A^{-1}を次のように表わす。

$$A^{-1} = B / |A| \qquad -(2.18)$$

ここで，

$$A = \begin{pmatrix} n & \sum_{i=1}^{n} H_i & \sum_{i=1}^{n} H_i^2 \\ \sum_{i=1}^{n} H_i & \sum_{i=1}^{n} H_i^2 & \sum_{i=1}^{n} H_i^3 \\ \sum_{i=1}^{n} H_i^2 & \sum_{i=1}^{n} H_i^3 & \sum_{i=1}^{n} H_i^4 \end{pmatrix}$$

また，B は A の余因子行列，$|A|$ は A の固有値。B の成分を b_{ij}で表わせば，

$$B = \begin{pmatrix} b_{11} & b_{12} & b_{13} \\ b_{21} & b_{22} & b_{23} \\ b_{31} & b_{32} & b_{33} \end{pmatrix} \qquad -(2.19)$$

となり，この場合 b_{12}と b_{21}，b_{13}と b_{31}，b_{23}と b_{32}が共役関係にあることが実証される。この表現を用いれば分散，共分散は以下のように表わされる。

$$\begin{aligned} \hat{V}(\hat{a}_0) &= (b_{11} / |A|)\cdot\hat{\sigma}^2 \\ \hat{V}(\hat{a}_1) &= (b_{22} / |A|)\cdot\hat{\sigma}^2 \\ \hat{V}(\hat{a}_2) &= (b_{33} / |A|)\cdot\hat{\sigma}^2 \\ \hat{Cov}(\hat{a}_0, \hat{a}_1) &= (b_{12} / |A|)\cdot\hat{\sigma}^2 \\ \hat{Cov}(\hat{a}_1, \hat{a}_2) &= (b_{23} / |A|)\cdot\hat{\sigma}^2 \\ \hat{Cov}(\hat{a}_2, \hat{a}_0) &= (b_{13} / |A|)\cdot\hat{\sigma}^2 \qquad -(2.20) \end{aligned}$$

ここで，

$$\begin{aligned} b_{11} &= \sum_{i=1}^{n} H_i^2 \sum_{i=1}^{n} H_i^4 - (\sum_{i=1}^{n} H_i^3)^2 \\ b_{22} &= n\sum_{i=1}^{n} H_i^4 - (\sum_{i=1}^{n} H_i^2)^2 \\ b_{33} &= n\sum_{i=1}^{n} H_i^2 - (\sum_{i=1}^{n} H_i)^2 \\ b_{12} &= \sum_{i=1}^{n} H_i^3 \sum_{i=1}^{n} H_i^2 - \sum_{i=1}^{n} H_i \sum_{i=1}^{n} H_i^4 \\ b_{23} &= \sum_{i=1}^{n} H_i \sum_{i=1}^{n} H_i^2 - n\sum_{i=1}^{n} H_i^3 \\ b_{13} &= \sum_{i=1}^{n} H_i \sum_{i=1}^{n} H_i^3 - (\sum_{i=1}^{n} H_i^2)^2 \\ |A| &= n\sum_{i=1}^{n} H_i^2 \sum_{i=1}^{n} H_i^4 + 2\sum_{i=1}^{n} H_i \sum_{i=1}^{n} H_i^2 \sum_{i=1}^{n} H_i^3 \\ &\quad - n(\sum_{i=1}^{n} H_i^3)^2 - (\sum_{i=1}^{n} H_i^2)^3 - \sum_{i=1}^{n} H_i^4 (\sum_{i=1}^{n} H_i)^2 \end{aligned}$$

これらの諸量を (2.15) 式に代入すれば，流量の分散推定値 $\hat{V}(Q)$が求まる。再度記すと，ある水位値 H_0に対して推定される流量 Q_0の分散 (推定値) は (2.15) 式から，

$$\begin{aligned} \hat{V}(Q_0) &= \hat{V}(\hat{a}_0) + H_0^2\hat{V}(\hat{a}_1) + H_0^4\hat{V}(\hat{a}_2) \\ &\quad + 2H_0\hat{Cov}(\hat{a}_0, \hat{a}_1) + 2H_0^3\hat{Cov}(\hat{a}_1, \hat{a}_2) \\ &\quad + 2H_0^2\hat{Cov}(\hat{a}_2, \hat{a}_0) \qquad -(2.21) \end{aligned}$$

となり，上式中の分散，共分散は (2.20) 式から求まる。このときの95%の信頼区間は，

$$(\hat{a}_0 + \hat{a}_1 H_0 + \hat{a}_2 H_0^2) \pm 1.96\sqrt{\hat{V}(Q_0)} \qquad -(2.22)$$

となる。

以上の作業で，直線回帰式および2次曲線回帰式の信頼区間が算出されるわけだが，これらを用いて実際の $H-Q$ 関係から得られる流量データの精度を評価していくことになる。

（補遺～1）正規方程式の一般形

多次元の回帰式の一般形は，以下の形となる。

$$\boldsymbol{y}=\hat{\boldsymbol{y}}+\boldsymbol{u}=X\boldsymbol{a}+\boldsymbol{u} \qquad \text{—①}$$

ここで，$\boldsymbol{y}$は従属変数のベクトル，$\hat{\boldsymbol{y}}$は$\boldsymbol{y}$の推定量，$\boldsymbol{u}$は誤差ベクトル，Xは説明変数のマトリックス，$\boldsymbol{a}$は回帰式の係数ベクトルで，従属変数y_iが説明変数x_iのべき乗形x_i, $x_i^2\cdots\cdots x_i^m$の線形結合で表わせるとする。各々の内訳を以下に示す。

$$\boldsymbol{y}=\begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix},\quad \hat{\boldsymbol{y}}=\begin{pmatrix} \hat{y}_1 \\ \hat{y}_2 \\ \vdots \\ \hat{y}_n \end{pmatrix},\quad \boldsymbol{u}=\begin{pmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{pmatrix}$$

$$X=\begin{pmatrix} 1 & x_1\cdots\cdots & x_1^{\ m} \\ 1 & x_2\cdots\cdots & x_2^{\ m} \\ \vdots & \ddots & \vdots \\ 1 & x_n\cdots\cdots & x_n^{\ m} \end{pmatrix},\quad \boldsymbol{a}=\begin{pmatrix} a_0 \\ a_1 \\ \vdots \\ a_m \end{pmatrix}$$

ただし，$(n\geqq m)$

このとき，以下の仮定を導入する。

（仮定1）　$E(\boldsymbol{u})=0$

これは，誤差項$\boldsymbol{u}$の期待値が0であることを示しており，通常とり入れられる仮定である。

（仮定2）　$E(\boldsymbol{u}\boldsymbol{u}')=\sigma^2 I_n$

ここで，$\boldsymbol{u}'$は$\boldsymbol{u}$の転置ベクトル，I_nはn次単位行列，σ^2は$V(\boldsymbol{y}-X\boldsymbol{a})$で示される誤差の分散である。すなわち，データ詳の誤差$\boldsymbol{u}$は一様の分散を持ち，データ間の共分散は0であることを示している。

今，$\boldsymbol{a}$の推定量$\hat{\boldsymbol{a}}$が求まったとき次式が成立する。

$$\boldsymbol{y}=X\hat{\boldsymbol{a}}+\boldsymbol{e} \qquad \text{—②}$$

ここで，$\boldsymbol{e}=\boldsymbol{y}-X\hat{\boldsymbol{a}}$（誤差ベクトル）

回帰式を得るための常とう手段として，最小二乗近似が用いられる。すなわち，上式の$\boldsymbol{e}$の二乗和を最小にしようとするもので，

$$\begin{aligned} \boldsymbol{e}'\boldsymbol{e} &= (\boldsymbol{y}-X\hat{\boldsymbol{a}})'(\boldsymbol{y}-X\hat{\boldsymbol{a}}) \\ &= (\boldsymbol{y}'-\hat{\boldsymbol{a}}'X')(\boldsymbol{y}-X\hat{\boldsymbol{a}}) \\ &= \boldsymbol{y}'\boldsymbol{y}-\boldsymbol{y}'X\hat{\boldsymbol{a}}-\hat{\boldsymbol{a}}'X'\boldsymbol{y}+\hat{\boldsymbol{a}}'X'X\hat{\boldsymbol{a}} \qquad \text{—③} \end{aligned}$$

ここで，$\boldsymbol{e}'$，$\boldsymbol{y}'$，$\hat{\boldsymbol{a}}$，X'は各々$\boldsymbol{e}$，$\boldsymbol{y}$，$\hat{\boldsymbol{a}}$，Xの転置ベクトル，転置マトリックスを表わす。

$$\frac{\partial}{\partial\hat{\boldsymbol{a}}}\left\{\boldsymbol{y}'X\hat{\boldsymbol{a}}\right\}=\frac{\partial}{\partial\hat{\boldsymbol{a}}}\left\{(X'\boldsymbol{y})'\hat{\boldsymbol{a}}\right\}=X'\boldsymbol{y} \qquad \text{—④}$$

$$\frac{\partial}{\partial\hat{\boldsymbol{a}}}\left\{\hat{\boldsymbol{a}}'X'\boldsymbol{y}\right\}=\frac{\partial}{\partial\hat{\boldsymbol{a}}}\left\{\hat{\boldsymbol{a}}'(X'\boldsymbol{y})\right\}=X'\boldsymbol{y} \qquad \text{—⑤}$$

$$\frac{\partial}{\partial\hat{\boldsymbol{a}}}\left\{\hat{\boldsymbol{a}}'X'X\hat{\boldsymbol{a}}\right\}=\frac{\partial}{\partial\hat{\boldsymbol{a}}}\left\{\hat{\boldsymbol{a}}'(X'X)\hat{\boldsymbol{a}}\right\}=2X'X\hat{\boldsymbol{a}} \qquad \text{—⑥}$$

などより，

$$\frac{\partial}{\partial\hat{\boldsymbol{a}}}\left\{\boldsymbol{e}'\boldsymbol{e}\right\}=-2X'\boldsymbol{y}+2X'X\hat{\boldsymbol{a}} \qquad \text{—⑦}$$

となり，ゆえに⑦式を**0**とおくと，次式を得る。

$$X'X\hat{\boldsymbol{a}}=X'\boldsymbol{y} \qquad \text{—⑧}$$

ここで，

$$X'X=\begin{pmatrix} n & \sum_{i=1}^{n}x_i & \cdots\cdots & \sum_{i=1}^{n}x_i^{\ m} \\ \sum_{i=1}^{n}x_i & \sum_{i=1}^{n}x_i^{\ 2} & \cdots\cdots & \sum_{i=1}^{n}x_i^{\ m+1} \\ \vdots & \vdots & \ddots & \vdots \\ \sum_{i=1}^{n}x_i^{\ m} & \sum_{i=1}^{n}x_i^{\ m+1} & \cdots\cdots & \sum_{i=1}^{n}x_i^{\ 2m} \end{pmatrix},\quad \hat{\boldsymbol{a}}=\begin{pmatrix} a_0 \\ a_1 \\ \vdots \\ a_m \end{pmatrix}$$

$$X'\boldsymbol{y}=\begin{pmatrix} \sum_{i=1}^{n}y_i \\ \sum_{i=1}^{n}x_i y_i \\ \vdots \\ \sum_{i=1}^{n}x_i^{\ m}y_i \end{pmatrix}$$

上式が一般的な正規方程式の形である。

（補遺～2）線形回帰式による推定値の分散と回帰係数の分散の関係について

ある確定値xから，回帰式によってyの値を推定することを考える。回帰式は次式のようなx，x^2，…，x^mの線形結合式とする。

$$y=a_0+a_1x+a_2x^2+\cdots+a_mx^m=\sum_{k=0}^{m}a_kx^k \qquad \text{—⑨}$$

データ群の分散によってyに推定誤差が生ずるが，これが回帰係数の統計的性質に帰因するとみる。すなわち，これが任意の確定値xからyを推定するとした場合，回帰係数a_k（$k=0$，1，…，m）が平均値$\bar{a}_k$のまわりに分散$V(a_k)$を持つとして，推定値yの分散は，

$$V(y)=\frac{1}{n}\sum_{i=1}^{n}(y_i-\bar{y})^2$$

$$=\frac{1}{n}\sum_{i=1}^{n}\Big\{(a_{0i}-\bar{a}_0)+x(a_{1i}-\bar{a}_1)+x^2(a_{2i}-\bar{a}_2)+\cdots+x^m(a_{mi}-\bar{a}_m)\Big\}^2$$

$$=V(a_0)+x^2V(a_1)+(x^2)^2V(a_2)+\cdots+(x^m)^2V(a_m)+2xCov(a_0,\ a_1)+2x\cdot x^2Cov(a_1,a_2)+\cdots+2x^{m-1}x^mCov(a_{m-1},\ a_m)$$

$$=\sum_{k=0}^{m}(x^k)^2V(a_k)+2\sum_{\substack{k=0\\(k\neq k')}}^{m}\sum_{k'=0}^{m}(x^k)(x^{k'})Cov(a_k,a_{k'}) \qquad —⑩$$

となる。これをさらに詳しくみると，⑨式より，

$$\frac{\partial y}{\partial a_k}=x^k \qquad —⑪$$

となることがわかるので，⑩式は，

$$V(y)=\sum_{k=0}^{m}\left(\frac{\partial y}{\partial a_k}\right)^2V(a_k)+2\sum_{\substack{k=0\\(k\neq k')}}^{m}\sum_{k'=0}^{m}\left(\frac{\partial y}{\partial a_k}\right)\cdot\left(\frac{\partial y}{\partial a_{k'}}\right)Cov(a_k,a_{k'}) \qquad —⑫$$

となる。なお，回帰係数 a_kの分散については補遺～3を参照のこと。

（補遺～3）回帰係数の分散について

⑧式の正規方程式を解いて，求められる回帰係数$\hat{\boldsymbol{a}}$の性質を調べてみる。⑧を$\hat{\boldsymbol{a}}$について解くと，

$$\hat{\boldsymbol{a}}=(X'X)^{-1}X'\boldsymbol{y} \qquad —⑬$$

ここで，$(X'X)^{-1}$は$X'X$の逆行列である。

なお，Xはm個の一次独立な変数の線形結合で表わされており，($m<n$（データ数)），よってランクはmである。また，これからX'のランクもmとなり，$X'X$はランクmの正則行列となって逆行列が存在することになる。次に，⑬式に①式を代入すると，

$$\hat{\boldsymbol{a}}=(X'X)^{-1}X'(X\boldsymbol{a}+\boldsymbol{u}) =\boldsymbol{a}+(X'X)^{-1}X'\boldsymbol{u} \qquad —⑭$$

が得られ，この期待値は，

$$E(\hat{\boldsymbol{a}})=\boldsymbol{a}+(X'X)^{-1}X'E(\boldsymbol{u}) =\boldsymbol{a} \qquad \because E(\boldsymbol{u})=0\ （仮定1） \qquad —⑮$$

となり，このことから$\hat{\boldsymbol{a}}$が$\boldsymbol{a}$の不偏推定量であることがわかる。

次に$\hat{\boldsymbol{a}}$の分散・共分散行列を求める。

$$(\hat{\boldsymbol{a}}-\boldsymbol{a})=(X'X)^{-1}X'\boldsymbol{u} \qquad —⑯$$

より，

$$V(\hat{\boldsymbol{a}})=E[(\hat{\boldsymbol{a}}-\boldsymbol{a})(\hat{\boldsymbol{a}}-\boldsymbol{a})']$$
$$=E[(X'X)^{-1}X'\boldsymbol{u}\boldsymbol{u}'X(X'X)^{-1}]$$
$$=(X'X)^{-1}X'E(\boldsymbol{u}\boldsymbol{u}')X(X'X)^{-1}$$
$$=(X'X)^{-1}X'\sigma^2I_nX(X'X)^{-1} \qquad \because E(\boldsymbol{u}\boldsymbol{u}')=\sigma^2I_n\ （仮定2） \qquad —⑰$$

となり，よって，

$$\underline{V(\hat{\boldsymbol{a}})=\sigma^2(X'X)^{-1}} \qquad —⑱$$

を得る。ちなみに右辺の$(X'X)$は，⑧式に示される正規方程式の係数行列である。このことから，最小二乗回帰係数の不偏推定量に関する分散・共分散行列は，正規方程式の係数行列の逆行列にデータの分散 σ^2をかけたものと等価であることが理解できる。なお σ^2の推定量 $\hat{\sigma}^2$を用いた場合は，回帰係数の分散，共分散も推定量として求まる。すなわち，

$$\hat{V}(\hat{\boldsymbol{a}})=\hat{\sigma}^2(X'X)^{-1} \qquad —⑲$$

となる。