AN AUTOMATIC SELECTION AND SOLVING METHOD FOR RATIONAL POLYNOMIAL COEFFICIENTS BASED ON NESTED REGRESSION

Tengfei LONG[a], Weili JIAO[b]*
[a] Center for Earth Observation and Digital Earth, Chinese Academy of Sciences
No.9 Dengzhuang South Road, Haidian District, Beijing 100094, China;
Tel: +86-010-82178188; Fax. +86-010- 82178009
E-mail: tflong@ceode.ac.cn
[b] Center for Earth Observation and Digital Earth, Chinese Academy of Sciences*
No.9 Dengzhuang South Road, Haidian District, Beijing 100094, China;
Tel: +86-010- 82178191; Fax. +86-010- 82178009
E-mail: wljiao@ceode.ac.cn

KEY WORDS: Rational function model; Nested regression; Goodness of fit; Coefficient of Determination; Variable selection; Ridge estimation

Abstract: As the development of high-resolution satellites, the rational function model (RFM) consisting of 78 rational polynomial coefficients (RPCs) is widely used to replace physical sensor models in photogrammetry and remote sensing. However, the correlation between the coefficients of RFM makes it difficult to solve the RPCs. In this paper, the problem of solving RPCs is converted into a problem of multiple linear regressions with serious multicollinearity, and a novel method based on nested regression is proposed to automatically select the proper RPCs. The significant coefficients of RFM are selected one by one according to the evaluation criteria of goodness of fit, while the redundancy coefficients are cast out, and the selected RPCs can be solved using ordinary least square method. Several satellite images including Quickbird P2AS, ALOS PRISM 1B2, SPOT5 HRG 1A and Landsat5 L2 are used in the tests, and the test results show that the proposed method could overcome the ill-condition and rank defect of the RFM. Generally speaking, geometric correcting accuracy using the new rational function model with no more than 20 selected line (row) RPCs is no worse than using the original model with 39 line (row) RPCs and ridge estimation (L-curve method), and the new model is hardly ill-conditioned. When the number of ground control points (GCPs) is less than 39, traditional RFM cannot be applied to geometric correction, while stable and accurate RPCs can also be obtained by utilizing the proposed method, and the geometric error of the result is less than 1 pixel.

INTRODUCTION

The rational function model (RFM) in remote sensing with 78 rational polynomial coefficients (RPCs) is a complete mathematical model, which approximately describes the physical imaging process in photogrammetry and remote sensing. Without knowing the position and orientation information of specific sensor, only plenty of ground control points (GCPs) are needed to solve all the unknown coefficients of the RFM and achieve high accuracies in the photogrammetric processing. In this sense, the RFM is suitable for high accuracy processing of different types of sensors, without disclosing the physical parameters of the sensor, so it is widely applied in photogrammetry and remote sensing. However, as the 78 RPCs of the RFM are strong correlated, stable and precision solutions of the RPCs are difficult or even impossible to achieve (Lin & Yuan, 2008; ZHU & JIAO, 2008).

Scholars did extensive researches on the RFM during the past ten years. OGC (Consortium et al., 1999) normalized the range of the image and object space coordinates of RFM to −1 to +1, and effectively enhanced the condition number of the normal equation matrix. Tao and Hu (Tao & Hu, 2001) comprehensively studied the RFM and suggested to overcome the ill-condition of the RFM using Tikhonov regularization and the L-curve method. Yuan and Lin (Yuan & Lin, 2008) compared the solving results of several RPCs solving methods including ridge trace method, L-curve method, empirical formula method, and generalized ridge estimate method, and verified the effectiveness of L-curve method. Moreover, Levenberg-Marquardt methods and singular value decomposition method are also applied to the solution of RPCs (Fraser et al., 2006). Among all these methods, the ridge estimation method (especially the L-curve method) is the most widely used, and it does well in overcoming the ill-condition of RFM. However, there are still some problems in RPCs solving that those existing methods cannot overcome, for example, ridge estimate is a biased estimate, many GCPs are required to solve the RPCs, and the RFM is lack of physical meanings.

Yuan and Cao (Yuan & Cao, 2011) proposed an optimized method for selecting RPCs, which enhances the stability of RPCs' solution. However, as the multiple correlations between variables are complex and no reliable evaluation method is available for the multiple correlations, removing some of the relevant variables often leads to increased interpretation error of the model, as well as discard of some information that should be retained (Geladi & Kowalski, 1986).

A nested regression method according to the evaluation criteria of goodness of fit is proposed. With the help of this method, the significant coefficients of RFM are gradually selected, while the redundancy coefficients are cast out, and the selected RPCs can be solved using ordinary least square method.

THE LINEARIZED FORM OF RFM

RFM is a pure mathematical model, which describes the object-to-image space transformation. In order to improve the numerical stability of the equations, image coordinates and object coordinates are both normalized to the range of -1.0 to 1.0. The RFM is given as (1):

$$\begin{cases} r = \dfrac{N_r(X,Y,Z)}{D_r(X,Y,Z)} \\ c = \dfrac{N_c(X,Y,Z)}{D_c(X,Y,Z)} \end{cases} \tag{1}$$

where r and c are normalized coordinates of image points in image space, while X, Y and Z are the normalized coordinates of ground points in object space. $N_r(X,Y,Z)$, $D_r(X,Y,Z)$, $N_c(X,Y,Z)$ and $D_c(X,Y,Z)$ are polynomials of X, Y and Z, whose coefficients are respectively $a_i, b_i, c_i, d_i (i = 0,1,\cdots,19)$. Taking $N_r(X,Y,Z)$ as an example, the polynomial is:

$$\begin{aligned} N_r(X,Y,Z) = {} & a_0 + a_1 Z + a_2 Y + a_3 X + a_4 ZY + a_5 ZX + a_6 YX + a_7 Z^2 \\ & + a_8 Y^2 + a_9 X^2 + a_{10} ZYX + a_{11} Z^2 Y + a_{12} Z^2 X + a_{13} ZY^2 \\ & + a_{14} Y^2 X + a_{15} ZX^2 + a_{16} YX^2 + a_{17} Z^3 + a_{18} Y^3 + a_{19} X^3 \end{aligned}$$

where $a_i, b_i, c_i, d_i (i = 0,1,\cdots,19)$ are the rational polynomial coefficients (RPCs). Generally speaking, b_0 and d_0 can be set as zero after reduction of the fraction, and they are supposed to be zero in this paper if there are no other special explains.

Although it is a nonlinear model, the RFM can be transformed into a linear model on all RPCs by a simple deformation, and the deformed linear model is shown as formula (2):

$$\begin{cases} N_r(X,Y,Z) - rD_r(X,Y,Z) = 0 \\ N_c(X,Y,Z) - cD_c(X,Y,Z) = 0 \end{cases} \tag{2}$$

As the formula (2) indicates the row RPCs and the column RPCs are independent of one another, it is possible to solve the row-RPCs and column RPCs separately. Taking the row RPCs for example, the first equation of formula (2) is rewritten as formula (3):

$$\begin{aligned} r = {} & a_0 + a_1 Z + a_2 Y + a_3 X + a_4 ZY + a_5 ZX + a_6 YX + a_7 Z^2 + a_8 Y^2 \\ & + a_9 X^2 + a_{10} ZYX + a_{11} Z^2 Y + a_{12} Z^2 X + a_{13} ZY^2 + a_{14} Y^2 X + a_{15} ZX^2 \\ & + a_{16} YX^2 + a_{17} Z^3 + a_{18} Y^3 + a_{19} X^3 - b_1 rZ - b_2 rY - b_3 rX - b_4 rZY \\ & - b_5 rZX - b_6 rYX - b_7 rZ^2 - b_8 rY^2 - b_9 rX^2 - b_{10} rZYX - b_{11} rZ^2 Y \\ & - b_{12} rZ^2 X - b_{13} rZY^2 - b_{14} rY^2 X - b_{15} rZX^2 - b_{16} rYX^2 - b_{17} rZ^3 \\ & - b_{18} rY^3 - b_{19} rX^3 \end{aligned} \tag{3}$$

When all the observation equations are collected, vector formula (4) is obtained:

$$\mathbf{r} = \mathbf{X}\beta_{\mathbf{r}} \tag{4}$$

where $\beta_{\mathbf{r}} = [a_0, a_1, \cdots, a_{19}, b_1, b_2, \cdots, b_{19}]^T$, $\mathbf{r} = \begin{bmatrix} r_1 \\ r_2 \\ \vdots \\ r_n \end{bmatrix}$, $\mathbf{X} = \begin{bmatrix} \mathbf{x_1} \\ \mathbf{x_2} \\ \vdots \\ \mathbf{x_n} \end{bmatrix}$, n is the number of observation equations,

$\mathbf{x_i} = \left[1, Z_i, Y_i, \cdots, X_i^3, -r_i Z_i, -r_i Y_i, \cdots, -r_i X_i^3\right]$, $i = 1, 2, \cdots, n$.

In a similar way, the column RPCs solving equations can be derived.

According to formula (4), the RPCs solving is essentially a multiple linear regression problem, and the estimated value of $\beta_{\mathbf{r}}$ may be obtained by ordinary least square method, shown as formula (5):

$$\hat{\beta}_{\mathbf{r}} = (\mathbf{X}^{\mathbf{T}}\mathbf{X})^{-1}\mathbf{X}^{\mathbf{T}}\mathbf{r} \tag{5}$$

where $\mathbf{X}^{\mathbf{T}}\mathbf{X}$ is the coefficient matrix of normal equation. Due to the strong correlation between the coefficients, $\mathbf{X}^{\mathbf{T}}\mathbf{X}$ is usually ill-posed. In this case, direct inverse of the matrix seldom produces accurate and stable results, and some ridge parameter is ordinarily brought in to reduce the ill-condition of the coefficient matrix.

GOODNESS OF FIT

The goodness of fit (Cameron & Windmeijer, 1997) of a statistical model describes how well the regression line approximates the observation data. Measures of goodness of fit typically summarize the discrepancy between observed values and the values expected under the model in question, including coefficient of determination and lack-of-fit sum of squares (Neter et al., 1996). In this paper, the coefficient of determination is used as the statistical measure of goodness of fit.

The coefficient of determination is the proportion of the regression sum of squares in total variation, and it is given as formula (6):

$$R^2 = \frac{\sum_{i=1}^{n} (\hat{y}_i - \overline{y})^2}{\sum_{i=1}^{n} (y_i - \overline{y})^2} \tag{6}$$

where n is the number of observation, y_i is the i_{th} observation value, $\overline{y}$ is the mean of all the observation values, $\hat{y}_i$ is the regression result of the i_{th} observation. The value of R^2 is in the range of 0 to 1. The greater is the value of R^2, the better does the regression line approximate the observation data, and the more suitable is the model (Hao, 2011).

OPTIMIZED SELECTION OF RPCS BASED ON NESTED REGRESSION

The concept of nested regression (Lin, 2008) is to divide the procedure of regression into several steps, and in each step, one variable that fits the objective vector best is selected from the variables set. When all the remained variables in the variables sets are not significant to the objective vector, the procedure of nested regression is finished.

Taking goodness of fit (coefficient of determination) as the evaluation criteria, modified nested regression method is used to optimal select the rational polynomial coefficients. And the procedure of the method is as following:

Step 1: Let the convergence thresholds $\mathrm{t}_1 = 0.5 / \mathrm{scale}$, and $\mathrm{t}_2 = 0.05 / \mathrm{scale}$, where scale is the magnification for coordinates normalization. Let $k = 1$, $\mathbf{r_k} = \mathbf{r}$, and no is the number of observations, and nv is the total number of RPCs. Then turn to step 2;

Step 2: Together with $\mathbf{r_k}$, $\mathbf{X_i}$ (where $\mathbf{X_i} \neq \mathbf{X_{(1)}} \sim \mathbf{X_{(k-1)}}$) is successively used to build the linear regression models:

$$\mathbf{r_k} = \beta_{k0}\mathbf{1} + \beta_{k1}\mathbf{X}_i + \varepsilon_\mathbf{k} \tag{7}$$

The coefficient of determination of each model is respectively calculated. Supposing the coefficient of determination of the model corresponding to $\mathbf{X_{(k)}}$ is the maximal one among all the coefficients of determination, formula (8) is derived from the model corresponding to $\mathbf{X_{(k)}}$ using ordinary least square regression:

$$\hat{\mathbf{r}}_\mathbf{k} = \hat{\beta}_{k0}\mathbf{1} + \hat{\beta}_{k1}\mathbf{X_{(k)}} \tag{8}$$

Then turn to step 3;

Step 3: If $k > no$ or $k > nv$, turn to step 4; otherwise the residual vector is $\mathbf{v_k} = \mathbf{r} - \sum_{m=1}^{k}\hat{\beta}_{mo}\mathbf{1} + \sum_{m=1}^{k}\hat{\beta}_{m1}\mathbf{X_{(m)}}$, and the root mean square error is $\sigma_k = \sqrt{\mathbf{v_k}^T\mathbf{v_k} / no}$. If $\sigma_k < t_1$ and $|\sigma_k - \sigma_{k-1}| < t_2$, turn to step 4; otherwise let $k = k+1$, and turn to step 2;

Step 4: optimized variables set $\{\mathbf{X_{(m)}} \mid m = 1, 2, \cdots, k\}$ is obtained, and ordinary least square method is used to regress the model(9), and the regression result is given as formula (10):

$$\mathbf{r} = \beta_0\mathbf{1} + \beta_1\mathbf{X_{(1)}} + \beta_2\mathbf{X_{(2)}} + \cdots + \beta_k\mathbf{X_{(k)}} + \varepsilon \tag{9}$$

$$\hat{\mathbf{r}} = \hat{\beta}_0\mathbf{1} + \hat{\beta}_1\mathbf{X_{(1)}} + \hat{\beta}_2\mathbf{X_{(2)}} + \cdots + \hat{\beta}_k\mathbf{X_{(k)}} \tag{10}$$

The coefficients those corresponding to $\mathbf{X_{(m)}}$ in $\beta_\mathbf{r}$ are set as $\hat{\beta}_m$ (where $m = 1, 2, \cdots, k$), and the other coefficients in $\beta_\mathbf{r}$ are set as 0. After that, the procedure of calculating the RPCs is finished.

EXPERIMENTAL RESULTS

In this section, two groups of tests are used to demonstrate that the proposed approach does indeed provide a practical way to get the optimal solution of RPCs.

1. Overcome the ill-condition of RFM

A scene of SPOT5 HRG 1A imagery in Beijing district of China, whose spatial resolution is 2.5 meters, is used in this test. 120 GCPs are well distributed in the range of the imagery, and 60 of them are evenly selected as observation data, while the other are used as check data. Three methods, including ordinary least square estimation, ridge estimation and nested estimation proposed in this paper, are applied to the calculation of the RPCs, and the check data are used to check the precision of the solutions. Table 1 shows the comparison of the results.

Table 1: Accuracy comparison of solution of three methods with 60 GCPs

	root-mean-square (RMS) error / pixel					
	least square regression		ridge regression		nested regression	
	row	column	row	column	row	column
GCPs	0.117	0.152	0.346	0.157	0.539	0.422
check points	2.64	0.675	1.25	0.87	0.685	0.79

According to table 1, although the least square estimation approximates the observation data well, the row RMS error of check data is more than 2 pixels. The reason is that the RPCs are not independent for the observation data and the model with 39 row (or column) RPCs is multicollinear and ill-conditioned, resulting the overfitting of observation data, as well as the instability of solution. The ridge parameter in ridge regression helps ameliorate the ill-condition of the model, and improves the stability of solution. However, the ridge regression does not cast out any unnecessary RPCs, and the row RMS error is still large. The nested regression gives stable results, and both the row and column RMS errors are less than 1 pixel. In fact, a number of unnecessary RPCs are cast out during the nested regression, and the simplified model is given as:

$$
\begin{aligned}
r &= \frac{-0.0037-0.0045Z-0.3075Y+1.1955X+0.0007Y-0.0001Z^2+0.0005Y^2-0.0001YX^2}{1.0+0.0027Y-0.0109X-0.0002ZY} \\
c &= \frac{0.0001Z-1.1343Y-0.2340X}{1.0-0.0002Y+0.0006X-0.0001YX-0.0001Y^2}
\end{aligned}
\tag{11}
$$

Comparing to formula (1), which contains 78 rational polynomial coefficients, formula (11) with only 18 coefficients is much more interpretable, and it shows clearly which physical quantities play important roles in the rational function model. Moreover, the row model is not necessarily the same as the column model, and the method of nested regression may adaptively build the suitable parsimonious models according to the observation data.

2. Rank Deficient Normal Equation

A scene of ALOS PRISM 1B2 imagery in Fujian district of China, whose spatial resolution is 2.5 meters, is used in this test, and 20 GCPs and 20 check points are prepared for this imagery. Three methods, including ordinary least square estimation, ridge estimation and nested estimation proposed in this paper, are respectively used to calculate the RPCs, and the check points are used to check the precision of the solutions. Table 2 shows the comparison of their results.

Table 2: accuracy comparison of solution of three methods with 20 GCPs

	root-mean-square (RMS) errors / pixel					
	least square regression		ridge regression		nested regression	
	row	column	row	column	row	column
GCPs	0.0	0.0	0.000095	0.000058	0.238	0.177
check points	11633.79	11561.21	841.72	851.81	0.644	0.144

When conventional methods (least square regression and ridge regression) are used to solve the row (or column) RPCs, there are 39 unknown variables, and no fewer than 39 GCPs are needed. As the rank of normal equation is deficient (fewer than 39), least square regression and ridge regression both fail to provide reliable solutions. By selecting significant RPCs from the full RFM (containing 78 RPCs), however, nested regression is able to provide reliable solutions of RPCs, and the RMS errors are both less than 1 pixel. The simplified RFM is given as:

$$
\begin{aligned}
r &= \frac{0.0001-0.0002Z-0.1410Y+1.0093X+0.0001YX}{1.0+0.0015Z^2X+0.0013ZY^2-0.0004ZX^2-0.0001X^3} \\
c &= -1.2808Y-0.3773X
\end{aligned}
\tag{12}
$$

According to formula (12), most of the coefficients in the full RFM are not necessary, and 20 GCPs are sufficient for the parsimonious model.

To verify the validity of the proposed method, some other tests are carried out on various satellite images including Quickbird P2AS, ALOS PRISM 1B2, SPOT5 HRG 1A and Landsat5 L2. Although the numbers of selected RPCs vary with the images, all the tests yield similar results, which indicate that the full RFM can be simplified into a new model of no more than 20 coefficients by means of optimized selection without reducing the accuracy of the full RFM. Moreover, the simplified models are much less likely ill-conditioned compared to the full RFM, and therefore more stable when applied to geometric correction.

CONCLUSIONS & RECOMMENDATIONS

The full RFM contains 78 RPCs, and the multicollinearity of the RPCs raises some critical problems, including the ill-condition and rank defect of the normal matrix. In this paper, the problem of solving rational polynomial coefficients is converted into a problem of multiple linear regressions with multicollinearity, and a complete methodology of automatically selecting and solving RPCs based on nested regression is proposed. By means of analysis of the observation data (GCPs), the proposed method casts out the unnecessary RPCs, and builds a simplified model. Both problems of ill-condition and rank defect are solved by the proposed method. Compare to the conventional methods (least square regression and ridge regression), the proposed nested regression provides a more stable and accurate solution when the observation data are sufficient. When the observation data are not plenty enough for the 78 unknown RPCs, reliable solutions are not available for the conventional methods. However, the proposed nested regression can provide reliable RPCs solution using fewer than 39 GCPs. In this sense, the proposed method would be quite useful when sufficient GCPs are not available. Moreover, the simplified RFM clearly shows which coefficients in the full RFM are significant for the individual imageries, and it helps to understand the imaging process.

ACKNOWLEDGEMENTS
The research has been supported by the National High Technology Research and Development Program of China (863 Program) under grant number 2006AA12Z118 and 2012BAH27B05.

REFERENCES:

Cameron, A. & Windmeijer, F. (1997), ‘An r-squared measure of goodness of fit for some common nonlinear regression models’, *Journal of Econometrics* **77**(2), 329–342.
Consortium, O.-O. et al. (1999), ‘The opengis abstract specification-topic 7: The earth imagery case’.
Fraser, C., Dial, G. & Grodecki, J. (2006), ‘Sensor orientation via rpcs’, *ISPRS journal of Photogrammetry and Remote Sensing* **60**(3), 182–194.
Geladi, P. & Kowalski, B. (1986), ‘Partial least-squares regression: a tutorial’, *Analytica chimica acta* **185**, 1–17.
Hao, P. (2011), *Econometrics*, Chengdu: Southwestern University of Finance and Economics Press.
Lin, F. (2008), ‘Solving multicollinearity in the process of fitting regression model using the nested estimate procedure’, *Quality & Quantity* **42**(3), 417–426.
Lin, X. & Yuan, X. (2008), ‘Improvement of the stability solving rational polynomial coefficients’, *the International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Beijing* **37**.
Neter, J., Wasserman, W., Kutner, M. & Li, W. (1996), *Applied linear statistical models*, number L, Irwin.
Tao, C. & Hu, Y. (2001), ‘A comprehensive study of the rational function model for photogrammetric processing’, *PE & RS- Photogrammetric Engineering & Remote Sensing* **67**(12), 1347–1357.
Yuan, X. & Cao, J. (2011), ‘An optimized method for selecting rational polynomial coefficients based on multicollinearity analysis’, *Geomatics and Information Science of Wuhan University* p. 06.
Yuan, X. & Lin, X. (2008), ‘A method for solving rational polynomial coefficients based on ridge estimation’, *Geomatics and Information Science of Wuhan University* **33**(11), 1130–1133.
ZHU, W. & JIAO, W. (2008), ‘Solving rational function model with genetic algorithm’, *Science Technology and Engineering* **8**, 97–100.