

УДК 520.88+521.95

Регрессионные модели в фотографической астрометрии

С. Г. Валеев, М. Г. Шамарин, И. А. Даутов, И. Е. Целищев

Для серии астронегативов анализируется структура редуцированных моделей, полученных при использовании нормальной схемы регрессионного анализа. Обоснована правомочность применения модели «плавающей» структуры. Предложено использование модели, состоящей из постоянной и переменной частей, изменяющихся по структуре от снимка к снимку.

THE REGRESSION MODELS IN PHOTOGRAPHIC ASTROMETRY, by Valeev S. G., Shamarin M. G., Dautov I. A., Tselishchev I. E. — For astrographic set of observations the structure of reductional models obtained by using the normal scheme of regression analysis is studied. The correctness of the application of «variable» structure model is justified and the model consisting of the constant and variable parts changing the structure from plate to plate is suggested.

Введение. Одна из основных проблем в теории и практике математической обработки результатов фотографических наблюдений — проблема выбора редуцированных моделей астрофотографий и повышения точности определения их параметров. Обычно в астрофотографии связь измеренных координат x , y с идеальными ξ , η представляется в виде полиномов

$$\begin{aligned}\xi - x &= ax + by + c + a'x^2 + b'xy + c'y^2 + \dots, \\ \eta - y &= dx + ey + f + d'x^2 + e'xy + f'y^2 + \dots,\end{aligned}\tag{1}$$

где коэффициенты $a, b, c, \dots; d, e, f, \dots$ определяются методом наименьших квадратов (МНК) и называются постоянными пластинки.

В данной работе анализируются результаты обработки фотографического ряда наблюдений звездных площадок в целях поиска оптимальных (по точности) редуцированных моделей для исходного математического описания типа (1). Задача решалась подходом регрессионного моделирования (РМ), основанного на регрессионном анализе (РА) [7]. Регрессионное моделирование выполнялось с помощью комплекса программных средств в виде автоматизированной системы обработки астрометрических баз данных, системное и функциональное наполнение которой описано в [3, 4]. Опыт применения подхода РМ в фотографической астрометрии отражен в некоторых публикациях [1, 2, 5]. Нами использована часть возможностей РМ, в полном объеме предусматривающего применение нормальной (стандартной) схемы, проверку соблюдения предположений стандартного РА и адаптацию стандартной схемы.

Постановка задачи и анализ результатов осуществлены С. Г. Валеевым, расчеты по моделированию — М. Г. Шамариным, исходные данные для моделирования в виде коэффициентов условных уравнений подготовлены И. А. Даутовым и И. Е. Целищевым.

Описание наблюдательного материала и исходных данных. Наблюдения выполнены в 1979—1981 гг. на широкоугольном астрографе Цейса ($D=40$ см, $F=2$ м), установленном на высокогорной станции Казанского университета [6]. Все астронегативы, кроме стандартной области Плеяд, получены в рамках работ по фотографированию южного неба. На каждой пластинке с определенной областью южного неба из-

мерено по 15—20 звезд, имеющих в каталоге Перт-70 [6]. Измерения (в которых принимали участие и авторы) проводились на координатометре «Аскорекорд». Средняя квадратичная ошибка измерения координат звезд составила 2—3 мкм, а для ярких звезд — до 7—8 мкм. В среднем на пластинку приходилась одна яркая звезда. Стандартная область Плеяд измерена автоматическим измерителем астронегативов (ИАН-2), который создан на основе прибора «Аскорекорд» в Астрономической обсерватории им. В. П. Энгельгардта. Средняя квадратичная ошибка измерения координат звезд оценивалась по повторяемости результатов и составила около 1 мкм.

Исходные данные для проведения РА представляют собой два массива: 1 — измеренные координаты звезд, исправленные за дисторсию объектива; 2 — идеальные координаты соответствующих звезд, полученные из каталога Перт-70 для астронегативов южного неба и специального каталога Якса [8] для области Плеяд. Каталогные сферические координаты приводились на видимые места с учетом влияния дифференциальной аберрации и рефракции. Нуль-пункт системы идеальных координат совмещался с оптическим центром пластинки. Средние квадратичные ошибки положений звезд в каталоге Перт-70 составляют $0.3''$ на эпоху 1980.0, в каталоге Якса равны $0.08''$ на эпоху около 1963.

Алгоритм и результаты применения нормальной схемы регрессионного анализа. Применение нормальной схемы означает: 1) принятие в качестве исходного описания зависимостей, в нашем случае вида (1), в которых $(\xi-x)$ и $(\eta-y)$ — отклики; x, y, x^2, xy, y^2 — регрессоры, в дальнейшем имеющие обозначения (1), (2), (1·1), (1·2), (2·2); a, b, c, a', b', c' — определяемые параметры; 2) применение в качестве метода оценивания МНК и статистический анализ уравнения; 3) принятие концепции о наличии множества конкурирующих моделей, насчитывающего в конкретном случае 2^5 элементов, и стратегии поиска наиболее информативного набора.

Представим каждое из уравнений (1) в виде модели РА

$$Y = X\beta + \varepsilon \quad (2)$$

где Y — отклик, случайный вектор размерности $(n \times 1)$, n — количество наблюдений; X — регрессионная матрица размерности $(n \times p)$, p — количество регрессоров; β — вектор оцениваемых параметров размерности $(p \times 1)$; ε — вектор ошибок размерности $(n \times 1)$.

Получаемые по МНК оценки $\hat{\beta}$ при соблюдении предположений РА, включая предположение о нормальности ошибок ε , оказываются наилучшими несмещенными оценками среди всех оценок, а не только в классе линейных несмещенных оценок. На каждом шаге построения модели $Y_i = \eta_i(X, \beta)$, $i = 1, l$, выполняется статистический анализ, включающий

расчет суммы квадратов остаточных уклонений $S = \sum_{i=1}^n V_i^2$ и функций от

S : стандартных ошибок наблюдений (ошибок единицы веса) σ и параметров $\hat{\beta}$, множественного коэффициента корреляции R , общего F -критерия и частных t -критериев.

Пусть задана исходная система признаков $x = \{x_1, x_2, \dots, x_p\}$ и некоторый критерий «ценности» совокупности признаков $S = S(x_{i_1}, x_{i_2}, \dots, x_{i_m})$, $\{i_k\}: 1 \leq i_1 < i_2 < \dots < i_m \leq p, k = 1, m$. В силу разного воздействия признаков на S и их взаимозависимости ценность признака будет определяться тем, с какой системой x он будет сочетаться. Необходимо указать такой набор m признаков ($m < p$), при котором значение критерия было бы оптимальным (для определенности — минимальным). Иначе, необходимо найти подсистему $x_m^* = \{x_{i_1}^*, x_{i_2}^*, \dots, x_{i_m}^*\}$, $x_m^* \subset x$, для которой $S(x_m^*) = \min_{x_m \subset x} S(x_m)$. При неизвестном m общее число подсистем равно 2^p .

Таблица 1. Структура оптимальных редукционных моделей и оценка их точности

Номер пластинки	Обозначение пластинки	Структура для П1	σ_{01}	σ_1	Структура для П2	σ_{02}	σ_2
1	A355	(1, 2)	0.971''	0.966''	(1, 2)	0.971''	1.158''
	(16)	(2)	0.902		(2)	0.902	1.152
2	A332	(1, 2)	0.550		(1, 2, 1·2)	0.530	0.584
	(20)	(2, 1)	0.401		(2, 1)	0.401	0.464
3	A307	(1, 2)	0.793		(1, 2)	0.793	0.907
	(17)	(2, 1)	0.906		(2, 1, 2·2)	0.817	0.947
4	A326	(1, 2)	0.623		(1, 2, 1·1)	0.543	0.546
	(24)	(2, 1)	0.852		(2, 1, 1·2)	0.602	0.628
5	A308	(1, 2)	0.450		(1, 2)	0.450	0.570
	(16)	(2, 1)	0.832		(2, 1)	0.832	0.906
6	Z2333	(1, 2)	0.580		(1, 2, 1·1)	0.567	0.624
	(18)	(2, 1)	0.690		(2, 1, 1·2, 2·2)	0.589	0.616
7	A325	(1, 2)	0.549		(1, 2, 1·2)	0.392	0.436
	(23)	(2, 1)	0.878		(2, 1)	0.878	0.996
8	A323	(1, 2)	0.862		(1, 2)	0.862	1.004
	(17)	(2, 1)	0.798		(2, 1, 1·2)	0.620	0.717
9	A321	(1, 2)	0.675		(1, 2, 1·1)	0.596	0.700
	(16)	(2, 1)	0.488		(2, 1, 1·2)	0.474	0.547
10	A284	(1)	0.459	0.469	(1, 1·2)	0.392	0.486
	(17)	(2)	0.722	0.736	(2)	0.722	0.828
11	A299	(1)	0.451	0.452	(1, 1·1)	0.358	0.384
	(22)	(2)	0.541	0.568	(2, 1·2)	0.538	0.613
12	A303	(1, 2)	0.874		(1, 2, 1·1, 2·2)	0.829	0.889
	(16)	(2, 1)	0.659		(2, 1, 2·2, 1·2)	0.644	0.701
13	A319	(1)	0.454	0.463	(1, 1·1)	0.323	0.364
	(21)	(2)	0.537	0.566	(2, 1·2)	0.384	0.432
14	A285	(1, 2)	0.700		(1, 1·1, 2)	0.619	0.703
	(20)	(2, 1)	0.463		(2, 1, 1·1, 1·2, 2·2)	0.359	0.359
15	A348	(1, 2)	0.412		(1, 2, 1·1)	0.395	0.426
	(18)	(2)	0.824		0.874	(2)	0.824
16	A272	(1, 2)	0.414		(1, 2)	0.414	0.507
	(16)	(2)	0.583		0.622	(2, 1·2)	0.552
17	A287	(1)	0.502	0.520	(1, 1·1)	0.425	0.508
	(18)	(2, 1)	0.550		(2, 1, 1·2, 2·2)	0.538	0.551
18	A350	(1, 2)	0.497		(1, 2, 1·1)	0.416	0.450
	(15)	(2, 1)	0.795		(2, 1)	0.795	0.970
19	A279	(1, 2)	0.809		(1, 2, 1·2)	0.708	0.742
	(20)	(2, 1)	0.680		(2, 1, 1·2)	0.658	0.720
20	Z2305	(1, 2)	0.814		(1, 2)	0.814	0.888
	(14)	(2, 1)	0.659		(2, 1, 1·1)	0.632	0.750
21	A300	(1, 2)	0.815		(1, 1·1)	0.586	0.644
	(20)	(2, 1)	0.926		(2, 1)	0.926	1.085
22	A309	(1)	0.658	0.671	(1, 1·1)	0.528	0.529
	(23)	(2)	0.835	0.864	(2)	0.835	0.987
23	A335	(1, 2)	0.551		(1, 1·1, 2)	0.491	0.557
	(17)	(2)	0.637		0.680	(2, 1·2)	0.533
24	A322	(1)	0.779	0.820	(1, 1·1)	0.770	0.934
	(17)	(2)	0.804	0.820	(2)	0.804	0.981
25	A304	(1, 2)	0.633		(1, 2)	0.633	0.753
	(19)	(2, 1)	0.715		(2, 1)	0.715	0.816
26	Z2331	(1, 2)	0.791		(1, 2)	0.791	1.013
	(15)	(2, 1)	0.923		(2, 1, 1·1)	0.907	0.999
27	A356	(1)	0.926	0.963	(1)	0.926	1.170
	(18)	(2)	0.728	0.765	(2, 1·2)	0.708	0.788
28	A290	(1)	0.856	0.916	(1, 1·1)	0.790	0.966
	(17)	(2)	0.501	0.533	(2)	0.501	0.614
29	A288	(1, 2)	0.981		(1, 2, 1·1)	0.907	1.052
	(16)	(2)	1.439		1.469	(2, 2·2)	1.276
30	A359	(1, 2)	0.672		(1, 2)	0.672	0.812
	(16)	(2, 1)	0.754		(2, 1, 1·1)	0.474	0.541
31	A343	(1, 2)	0.458		(1, 2·2, 2, 1·2, 1·1)	0.378	0.378
	(15)	(2, 1)	0.581		(2, 1, 1·2)	0.555	0.655
32	A289	(1, 2)	0.418		(1, 2)	0.418	0.489
	(17)	(2, 1)	0.472		(2, 1·2, 1)	0.400	0.442
33	A302	(1, 2)	1.037		(1, 2, 2·2)	1.020	1.098
	(18)	(2, 1)	0.874		(2, 1)	0.874	1.032

Номер пластинки	Обозначение пластинки	Структура для П 1	σ_{01}	σ_1	Структура для П 2	σ_{02}	σ_2
34	A317	(1, 2)	0.569"		(1, 2, 1·1)	0.551"	0.618"
	(21)	(2)	0.503	0.503"	(2)	0.503	0.570
35	A344	(1)	1.008	1.075	(1)	1.008	1.184
	(22)	(2)	0.739	0.818	(2)	0.739	0.895
36	A346	(1, 2)	0.755		(1, 2, 1·1, 1·2)	0.403	0.428
	(17)	(2, 1)	0.880		(2, 1, 1·2, 2·2)	0.478	0.519
37	A316	(1, 2)	0.656		(1, 2)	0.656	0.784
	(18)	(2, 1)	0.885		(2, 1, 2·2)	0.776	0.888
38	A318	(1, 2)	0.446		(1, 2, 1·1)	0.432	0.497
	(18)	(2, 1)	0.880		(2, 1)	0.880	1.027
39	Z2330	(1, 2)	0.558		(1, 2, 1·2)	0.454	0.550
	(14)	(2, 1)	0.724		(2, 1)	0.724	0.926
40	A283	(1, 2)	0.283		(1, 2)	0.283	0.334
	(17)	(2, 1)	0.640		(2, 1, 1·2)	0.603	0.699
41	Z2306	(1, 2)	0.475		(1, 2, 2·2)	0.419	0.467
	(13)	(2, 1)	0.555		(2, 1)	0.555	0.690
42	A315	(1, 2)	0.612		(1, 2, 1·2)	0.641	0.762
	(16)	(2, 1)	0.718		(2, 1·1, 1·2, 2·2, 1)	0.382	0.382
43	A320	(1, 2)	0.595		(1, 2)	0.595	0.752
	(15)	(2, 1)	0.584		(2, 1, 1·2)	0.527	0.586
44	A332	(1)	0.553	0.559	(1, 2, 1·2)	0.550	0.619
	(19)	(2, 1)	0.420		(2, 1)	0.420	0.489
45	Z2307	(1, 2)	0.490		(1, 2, 2·2)	0.388	0.440
	(16)	(2, 1)	0.567		(2, 1, 1·1, 1·2·2)	0.555	0.608
46	A293	(1, 2)	0.802		(1, 2, 1·1)	0.776	0.899
	(18)	(2)	0.787	0.837	(2)	0.787	0.989
47	A296	(1, 2)	0.662		(1, 2, 1·1)	0.584	0.632
	(22)	(2, 1)	0.818		(2, 1, 1·2, 2·2)	0.460	0.488
48	A313	(1, 2)	0.700		(1, 2)	0.700	0.808
	(19)	(2, 1)	0.613		(2, 1·2, 1·1)	0.534	0.571
49	A333	(1, 2)	0.430		(1, 2, 1·1, 1·2)	0.319	0.328
	(17)	(2, 1)	0.561		(2, 1)	0.561	0.642

Примечание. Во второй графе в скобках — число использованных опорных звезд; σ_{01} и σ_{02} — средняя квадратичная ошибка для соответствующей модели; σ_1 и σ_2 — средняя квадратичная ошибка для исходного полинома первой и второй степени, при совпадении оптимальной модели с исходной ошибка не указывается.

Метод поиска, примененный в работе, соответствует методу полного перебора [7].

Результаты применения нормальной схемы РА при обработке 49 астрофотографий приведены в табл. 1, в которой содержатся описание оптимальных по σ структур регрессионных моделей (т. е. номера введенных в модель регрессоров, при совокупности которых фиксируется σ_{\min}) и информация о точности в виде стандартной ошибки оценки σ (ошибки единицы веса в секундах дуги) для исходной и оптимальной моделей, вычисляемой по формуле $\sigma = \sqrt{\frac{\sum V_i^2}{n-1}}$, где V_i — уклонения; n — количество опорных звезд.

На основе выражений (1) строились исходные модели в виде полиномов первой и второй степени по измеренным координатам x и y , причем во всех моделях предполагается наличие членов c и f .

Рассмотрим результаты моделирования при исходной редуцированной модели в виде полинома первой степени (П1). По координатам ξ и η соответственно в 80 и 69 % случаев фиксируется набор регрессоров (1, 2) и (2, 1), что отвечает исходному описанию. Оптимальные описания, содержащие регрессоры (1) или (2), наблюдаются соответственно в 20 и 31 % случаев, обеспечивая повышение внутренней точности по σ в среднем на 3 и 5 % по координатам ξ и η .

Таким образом, даже при малом числе конкурирующих моделей, которое порождается исходным набором (x, y) , обнаруживаются тен-

денции к непостоянству структуры (состава) редуциционной модели для астрофотографий, полученных на одном и том же астрографе, к уменьшению размерности оптимальной модели по сравнению с исходной, т. е. соответственно в 20 и 31 % всех случаев регрессоры (2) для ξ и (1) для η оказались малоинформативными или несущественными вообще.

Анализ результатов моделирования при исходной редуциционной модели в виде полинома второй степени (П2) приводит к следующим заключениям. По координатам ξ и η исходное описание (1, 2, 1·1, 1·2, 2·2) оказывается оптимальным по критерию σ только соответственно в 6 и 4 % случаев. Оптимальные описания разного состава, не совпадающие с исходным, обнаруживаются по ξ и η в 94 и 96 % случаев, обеспечивая в среднем повышение внутренней точности по σ на 12 %, при включении меньшего количества членов по сравнению с исходным полиномом второй степени.

Таким образом, при исходном описании П2 оптимальные редуцированные модели непостоянны по структуре для астрофотографий, полученных на одном телескопе, содержат меньшее число членов по сравнению с исходным описанием и обеспечивают повышение внутренней точности по σ на 12 %, что представляет практический интерес.

Отметим причины, приводящие к различию модели по структуре от пластинки к пластинке. Назовем такую модель моделью «плавающей» структуры. Косвенные соображения и расчеты, проведенные с исключением отдельных звезд, дающих при обработке большие отклонения, свидетельствуют о том, что в измеренных и каталожных координатах звезд есть систематические ошибки разной природы, изменяющиеся от пластинки к пластинке и обуславливающие тем самым включение в модель различных членов П2. Исключение из обработки только одной из опорных звезд, для которых обнаруживаются большие отклонения, может привести к отбрасыванию последнего члена в модели или к понижению по абсолютной величине соответствующей t -статистики, характеризующей степень существенности регрессора.

Несмотря на то, что модели «плавающей» структуры обеспечивают лучшую привязку к системе опорных звезд астронегатива, они могут показаться (при отсутствии соответствующих программных средств) неудобными в обращении. Описанная автоматизированная система снимает эту проблему. При необходимости можно сформулировать редуциционные модели постоянной структуры по ξ и η , которые должны включать регрессоры, чаще всего повторяющиеся в «плавающих» моделях. В нашем случае (табл. 1) в моделях по ξ регрессоры (1), (2), (1·1) повторяются соответственно 49, 39 и 23 раза, а в моделях по η регрессоры (1), (2), (1·2) — 33, 49, 21 раз. Следовательно, обработанной фотографической серии соответствуют модели постоянной структуры

$$\xi - x = ax + by + c + a'x^2, \quad \eta - y = dx + ey + f + e'xy.$$

На наш взгляд, для полного использования возможностей этих двух типов моделей необходим переход к комплексной модели, состоящей из принудительно включенной постоянной (для учета геометрии перехода и стабильных систематических эффектов) и переменной (учитывающей изменяющиеся от пластинки к пластинке систематические ошибки) частей.

В заключение представим кратко результаты обработки двух снимков стандартной области Плеяд, содержащих практически одни и те же звезды, но полученные в разные дни одного месяца с несколько различающимися экспозициями. Каждый из снимков обработан в двух вариантах: 1 — исходная модель содержит координаты x (1), y (2); 2 — x (1), y (2), фотографическую звездную величину m (5), радиус изображения звезды R (6). Основные результаты приведены в табл. 2, которая по построению аналогична табл. 1.

Согласно данным табл. 2, модели снимка Z1235 обладают лучшей внутренней точностью. Непосредственные результаты регрессионного моделирования: 1. Оптимальной моделью для первого варианта является П1. Отсутствие квадратичных членов по измеренным координатам легко объяснить малостью звездной области фотографирования, составляющей $1.5 \times 1.5^\circ$, по сравнению с $4.5 \times 4.5^\circ$ для рассмотренных ранее снимков. 2. Оптимальные модели для П1 второго варианта включают дополнительные факторы m или R по одной из координат, при этом точность повышается. Редукционные модели по обоим снимкам не совпадают по структуре, что, видимо, обусловлено различной продолжительностью экспозиции (4.5 мин для Z1235 и 3 мин для Z1229). 3. Оптимальные модели для П2 второго варианта включают по ξ и η дополнительные члены в виде взаимодействия m и R , при этом точность повышается как по отношению к полной модели П2, так и к оптимальной по П1. Как и в предыдущем случае, модели по обоим снимкам не совпадают по структуре: наблюдается общая тенденция к убыванию средних квадратичных ошибок единицы веса.

Таблица 2. Структура оптимальных моделей для стандартной области Плеяд и оценка их точности

Вариант	Обозначение пластики	Структура для П1	σ_{01}	σ_1	Структура для П2	σ_{02}	σ_2
1	Z1229	(1, 2)	0.372"	0.372"	(1, 2)	0.372"	0.368"
	(49)	(2, 1)	0.390	0.390	(2, 1)	0.390	0.394
1	Z1235	(1, 2)	0.299	0.299	(1, 2)	0.299	0.295
	(51)	(2, 1)	0.298	0.298	(2, 1)	0.298	0.312
2	Z1229	(1, 2)	0.372	0.377	(1, 2, 1-5, 2-5)	0.284	0.309
	(49)	(2, 1, 6)	0.350	0.357	(2, 1, 2-6)	0.255	0.283
2	Z1235	(1, 2, 5)	0.265	0.270	(1, 2, 1-5, 1-1)	0.201	0.220
	(51)	(2, 1)	0.298	0.306	(2, 1, 2-5)	0.231	0.238

Таким образом, применение в фотографической астрометрии подхода РМ во всех случаях представляется полезным. Если исследователь считает, что предложенная им модель является причиной обусловленной и единственно возможной, то применение РМ позволяет установить, так ли это на самом деле. Для этого достаточно исследовать отклонения $(y_i - \hat{y}_i)$, где \hat{y}_i — прогнозируемые значения, и выяснить, не содержат ли они систематических эффектов, которые могут быть описаны каким-либо аппроксимирующим выражением с заметным коэффициентом множественной корреляции R .

1. Валеев С. Г. О применении метода множественной регрессии в астрометрии // Задачи современной астрометрии в создании всероссийской системы координат.— Ташкент: Фан, 1981.— С. 125—129.
2. Валеев С. Г. // Изв. вузов. Геодезия и аэрофотоъемка.— 1982.— № 2.— С. 121—122.
3. Валеев С. Г. Функциональное содержание пакета прикладных программ по автоматизации обработки астрометрических баз данных // Методы и программное обеспечение обработки информации и прикладного статистического анализа данных на ЭВМ.— Минск: Изд-во Белорус. ун-та, 1985.— С. 213—214.
4. Валеев С. Г., Мадер В. Г., Толлок Б. Б., Шамарин М. Г. Математическое обеспечение автоматизированной системы контроля качества продукции комплексной АСУ ТП // Математические вопросы проектирования территориальных АСУ.— Кемерово: Изд-во Кемеров. ун-та, 1983.— С. 58—65.
5. Валеев С. Г., Чекушева Т. Д. Трансформация координат методом множественной регрессии // Молодые ученые и специалисты Кемеровской области — народному хозяйству.— Кемерово: Науч.-техн. о-во, 1977.— С. 97—98.
6. Даутов И. А., Жиков Г. В., Маниров Т. К. и др. // Тр. Казан. гор. астроном. обсерватории.— 1983.— № 48.— С. 117—122.
7. Себер Д. Линейный регрессионный анализ.— М.: Мир, 1980.— 456 с.
8. Jaks W. // Publ. Inst. Geophys. Pol. Acad. Sci.— 1977.— N 108.— P. 95—105.