

マルコフ連鎖に基づく打者評価モデル

穴太克則

南山大学経営学部情報管理学科

Evaluation Model for Batter based on Markov Chain.

Katsunori Ano

Faculty of Business, Nanzan University

Abstract. We propose an improved Offensive Earned-Run Average(OERA) value which is based on a stationary Markov Chain and is a well-known index to evaluate a contribution of batter in baseball game. However an original OERA value doesn't consider steal's effect. It seems to be unrealistic. Proposed value in this paper includes the effect. Some examples are demonstrated.

1. はじめに

野球における打者貢献度の評価方法は様々に考えられている。本塁打数、打率、打点などもそのひとつの指標と言える。しかし、これだけでは、本塁打は多いが打率の低い打者と本塁打は少ないが打率の高い打者とではどちらの貢献度が高いかという判断は難しい。また、チャンスに強く打点が多いが出塁数が少ない打者と、チャンスに弱く打点が少ないが出塁数が多い打者の比較なども同様である。これらの短所を補うモデルとしては Cover and Keilers[1] の OERA(Offensive Earned-Run Average) 値が有名である。次章で詳しく述べるが、この OERA 値は 1 つの吸収状態をもつマルコフ連鎖を基にして打者の貢献度を数値化するシンプルなモデルである。しかし残念ながら OERA のモデルでは盗塁が加味されていない。四死球を選び盗塁し 2 塁まで進塁した場合は二塁打と同じ価値を持つ場合がある。例えば '95 年度パ・リーグでは、49 盗塁したイチローの OERA 値は 9.7469(.342, 25 本塁打, 86 四死球), 6 盗塁であるオマリー(.302, 31 本塁打, 98 四死球) の OERA 値は 10.0390 となり、オマリーがイチローを上回る。攻走守の三拍子そろった打者がよいと言われているが、オリジナルな OERA 値では「攻」のみしか評価されず、「走」の能力が秀でている選手の貢献は無視されてしまう。本論文では、「走」の要素である盗塁を評価に加えてマルコフ連鎖の推移確率行列を改良した OERA 値(TOERA 値と称す)を提案、解説し、OERA 値と TOERA 値の計算例を紹介する。提案される TOERA 値では、イチローが 9.87459 (49 盗塁, 9 盗塁死), オマリーが 9.74721(6 盗塁, 6 盗塁死) となり、イチローがオマリーを上回る。TOERA 値はシンプルなモデルながらより現実的であり、この値を基に打者年棒評価への適用等の実用性に富むと思え、ここに報告する。

2. TOERA 値

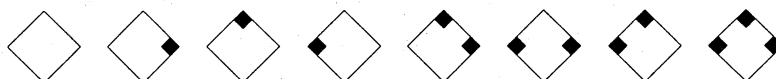
打者評価法としての OERA 値は、一人の特定の打者が常に打席に立ち 9 回まで攻撃したと仮定して何点の得点が期待できるかを基準とする。まず、凡打、単打、二塁打、三塁打、本塁打、四死球、盗塁に対してアウトカウントとランナーの状態がどのように推移するかの規則を明確にしておく。

規則

1. 犠打はすべて計算されない。
2. エラーはアウトとして計算される。
3. アウトによってランナーは進塁しない。
4. 単打は一塁ランナーを三塁へ進塁させ、二塁ランナーと三塁ランナーをホームへ生還させる。
5. 二塁打と三塁打は一塁ランナー、二塁ランナー、三塁ランナー すべてをホームへ生還させる。
6. ダブルプレーはないとする。
7. 盗塁は安打の後のみに試みることができるとし、その試みは 1 回とする。すなわち二盗と三盗、三盗と本盗というように 2 回続けては盗塁しない。

次に起こり得る状態を下図のように定義する。

状態



ノーアウト	1	2	3	4	5	6	7	8
ワンアウト	9	10	11	12	13	14	15	16
ツーアウト	17	18	19	20	21	22	23	24
スリーアウト	0							

状態空間を S とすると $S = \{0, 1, \dots, 24\}$ であり、スリーアウトを 0、ノーアウトランナーなしを 1、ノーアウトランナー 1 塁を 2, ..., ツーアウト満塁を 24 とする。スリーアウトは吸収状態となる。打撃と盗塁という行動により状態が推移する。打撃と盗塁に関する確率を以下に定義する。

打撃と盗塁に関する確率

$$p_0 = P_r(\text{凡打}) = \frac{\text{凡打数}}{\text{打数} + \text{四死球数}}$$

$$p_B = P_r(\text{四死球}) = \frac{\text{四死球数}}{\text{打数} + \text{四死球数}}$$

$$p_1 = P_r(\text{単打して二盗失敗}) = \frac{\text{単打数}}{\text{打数} + \text{四死球数}} \cdot \frac{\text{単打のときの二盗試行数}}{\text{単打数}} \times \frac{\text{単打のときの二盗失敗数}}{\text{単打のときの二盗試行数}}$$

$$p_2 = P_r(\text{単打して二盗しない}) = \frac{\text{単打数}}{\text{打数} + \text{四死球数}} \cdot \frac{\text{単打のときの二盗不試行数}}{\text{単打数}}$$

$$p_3 = P_r(\text{単打して二盗成功}) = \frac{\text{単打数}}{\text{打数} + \text{四死球数}} \cdot \frac{\text{単打のときの二盗試行数}}{\text{単打数}} \times \frac{\text{単打のときの二盗成功数}}{\text{単打のときの二盗試行数}}$$

$$p_4 = P_r(\text{二塁打して三盗失敗}) = \frac{\text{二塁打数}}{\text{打数} + \text{四死球数}} \cdot \frac{\text{二塁打のときの三盗失敗数}}{\text{二塁打のときの三盗試行数}}$$

$$p_5 = P_r(\text{二塁打して三盗しない}) = \frac{\text{二塁打数}}{\text{打数} + \text{四死球数}} \cdot \frac{\text{二塁打のときの未盗塁数}}{\text{二塁打数}}$$

$$p_6 = P_r(\text{二塁打して三盗成功}) = \frac{\text{二塁打数}}{\text{打数} + \text{四死球数}} \cdot \frac{\text{二塁打のときの三盗成功数}}{\text{二塁打のときの三盗試行数}}$$

$$p_7 = P_r(\text{三塁打して本盗失敗}) = \frac{\text{三塁打数}}{\text{打数} + \text{四死球数}} \cdot \frac{\text{三塁打のときの本盗失敗数}}{\text{三塁打のときの本盗試行数}}$$

$$p_8 = P_r(\text{三塁打して本盗しない}) = \frac{\text{三塁打数}}{\text{打数} + \text{四死球数}} \cdot \frac{\text{三塁打のときの未盗塁数}}{\text{三塁打数}}$$

$$p_9 = P_r(\text{三塁打して本盗成功}) = \frac{\text{三塁打数}}{\text{打数} + \text{四死球数}} \cdot \frac{\text{三塁打のときの本盗成功数}}{\text{三塁打のときの本盗試行数}}$$

$$p_{10} = P_r(\text{本塁打}) = \frac{\text{本塁打数}}{\text{打数} + \text{四死球数}}$$

推移確率行列

推移確率行列 $P = (P_{ij}) = p(j|i)$, $i, j = 0, 1, \dots, 24$ は次のようになる。

$$P = \begin{bmatrix} 1 & 0 \\ T & Q \end{bmatrix}$$

規則に従えば, T, Q は次のように表わされる。

$$T = \begin{matrix} & & 0 \\ 1 & & \\ \vdots & & 0 \\ 16 & & \\ 17 & p_0 + p_1 + p_7 & \\ \vdots & & \vdots \\ 24 & p_0 + p_1 + p_7 & \end{matrix} \quad , \quad Q = \begin{matrix} & 1 \dots 8 \ 9 \dots 16 \ 17 \dots 24 \\ 1 & Q_{11} \ Q_{12} \ Q_{13} \\ \vdots & \\ 8 & \\ 9 & \\ \vdots & Q_{21} \ Q_{22} \ Q_{23} \\ 16 & \\ 17 & \\ \vdots & \\ 24 & Q_{31} \ Q_{32} \ Q_{33} \end{matrix} .$$

ただし,

$$Q_{11} = \begin{bmatrix} p_9 + p_{10} & p_B + p_2 & p_3 + p_5 & p_6 + p_8 & 0 & 0 & 0 & 0 \\ p_9 + p_{10} & 0 & p_5 & p_6 + p_8 & p_B & p_2 & p_3 & 0 \\ p_9 + p_{10} & p_2 & p_3 + p_5 & p_6 + p_8 & p_B & 0 & 0 & 0 \\ p_9 + p_{10} & p_2 & p_3 + p_5 & p_6 + p_8 & 0 & p_B & 0 & 0 \\ p_9 + p_{10} & 0 & p_5 & p_6 + p_8 & 0 & p_2 & p_3 & p_B \\ p_9 + p_{10} & 0 & p_5 & p_6 + p_8 & 0 & p_2 & p_3 & p_B \\ p_9 + p_{10} & p_2 & p_3 + p_5 & p_6 + p_8 & 0 & 0 & 0 & p_B \\ p_9 + p_{10} & 0 & p_5 & p_6 + p_8 & 0 & p_2 & p_3 & p_B \end{bmatrix} ,$$

$$Q_{12} = \begin{bmatrix} p_0 + p_1 + p_4 + p_7 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ p_4 + p_7 & p_0 & 0 & p_1 & 0 & 0 & 0 & 0 \\ p_1 + p_4 + p_7 & 0 & p_0 & 0 & 0 & 0 & 0 & 0 \\ p_1 + p_4 + p_7 & 0 & 0 & p_0 & 0 & 0 & 0 & 0 \\ p_4 + p_7 & 0 & 0 & p_1 & p_0 & 0 & 0 & 0 \\ p_4 + p_7 & 0 & 0 & p_1 & 0 & p_0 & 0 & 0 \\ p_1 + p_4 + p_7 & 0 & 0 & 0 & 0 & 0 & p_0 & 0 \\ p_4 + p_7 & 0 & 0 & p_1 & 0 & 0 & 0 & p_0 \end{bmatrix} ,$$

$$Q_{13} = Q_{21} = Q_{31} = Q_{32} = 0 \quad (8 \times 8 \text{の零行列}),$$

$$Q_{11} = Q_{22} = Q_{33}, \quad Q_{12} = Q_{23}.$$

例えば, 状態 1(ノーアウトランナーなし) から状態 4(ノーアウトランナー三塁) になるのは, 二塁打を打って盗塁に成功するか, 三塁打を打って本盗しないかの どちらかの場合に限られるので, 状態 1 から状態 4 への推移確率は $p(4|1) = p_6 + p_8$ となる. 同様にして他の推移確率が計算できる.

TOERA 値の計算

$E(i)$ を状態 i から始まるイニングの期待得点とする. 各イニングは状態 1 より始まるので, 1 イニングの期待得点は $E(1)$ となる. これより, $E(1)$ の値を計算できれば, それを

9 イニング倍することにより TOERA 値が求まる. すなわち, TOERA 値は

$$(1.1) \quad TOERA = 9E(1),$$

により与えられる.

$E(i)$ の計算方法を述べる. 状態 i から状態 j に移ったときの得点を $R(j, i)$ とする. このときマルコフ連鎖の First Step Analysis (例えば, [5] 参照) より,

$$E(i) = \sum_{j=1}^{24} p(j|i) \{R(j, i) + E(j)\}, \quad i = 1, \dots, 24.$$

が成立する. $\sum_{j=1}^{24} p(j|i)R(j, i)$ は 状態 i における期待得点を表すので, この期待得点を以下のように $R(i)$ とおく.

$$R(i) = \sum_{j=1}^{24} p(j|i)R(j, i), \quad i = 1, \dots, 24.$$

ここで, R, E を

$$R = \begin{bmatrix} R(1) \\ \vdots \\ R(24) \end{bmatrix}, \quad E = \begin{bmatrix} E(1) \\ \vdots \\ E(24) \end{bmatrix},$$

のようにベクトルで表せば, (1.2) 式は

$$E = QE + R$$

となり, E は

$$(1.2) \quad E = (I - Q)^{-1}R$$

により求まる. 状態 1 から始まる 1 イニングあたりの期待得点はベクトル E の要素 $E(1)$ となり, これより TOERA 値が得られる. (1.2) 式は次のように考えることができる. 非吸収状態間の推移確率行列 Q に対して,

$$(I - Q)^{-1} = I + Q^2 + Q^3 + \dots$$

なる関係が成立し, $(I - Q)^{-1}$ は吸収マルコフ連鎖の基本行列と呼ばれる. この基本行列の i, j 要素は, 状態 i より始めて, 状態 j を通過する期待回数を表す. 野球は状態 1 より始まるので, 基本行列の $1j$ 要素である通過期待回数に状態 j における期待得点値 R をかければ, 1 イニングの期待得点値が得られる. 最後に状態 i における期待得点 $R(i)$ は, 規則に従えば次のように表される.

$$R = \begin{bmatrix} R(1) \\ R(2) \\ \vdots \\ R(24) \end{bmatrix} = \begin{matrix} 1 \\ \vdots \\ 8 \\ 9 \\ \vdots \\ 16 \\ 17 \\ \vdots \\ 24 \end{matrix} \begin{bmatrix} R_1 \\ R_2 \\ R_3 \end{bmatrix},$$

$$\begin{aligned}
 R_1 = R_2 &= \begin{bmatrix} p_9 + p_{10} \\ 2p_9 + 2p_{10} + p_4 + p_5 + p_6 + p_7 + p_8 \\ 2p_9 + 2p_{10} + p_1 + p_2 + p_3 + p_4 + p_5 + p_6 + p_7 + p_8 \\ 2p_9 + 2p_{10} + p_1 + p_2 + p_3 + p_4 + p_5 + p_6 + p_7 + p_8 \\ 3p_9 + 3p_{10} + 2p_4 + 2p_5 + 2p_6 + 2p_7 + 2p_8 + p_1 + p_2 + p_3 \\ 3p_9 + 3p_{10} + 2p_4 + 2p_5 + 2p_6 + 2p_7 + 2p_8 + p_1 + p_2 + p_3 \\ 3p_9 + 3p_{10} + 2p_1 + 2p_2 + 2p_3 + 2p_4 + 2p_5 + 2p_6 + 2p_7 + 2p_8 \\ 4p_9 + 4p_{10} + 3p_1 + 3p_4 + 3p_5 + 3p_6 + 3p_7 + 3p_8 + 2p_2 + 2p_3 + p_B \end{bmatrix}, \\
 R_3 &= R_1 + \begin{bmatrix} 0 \\ p_4 + p_7 \\ p_1 + p_4 + p_7 \\ p_1 + p_4 + p_7 \\ p_1 + 2p_4 + 2p_7 \\ p_1 + 2p_4 + 2p_7 \\ 2p_1 + 2p_4 + 2p_7 \\ 2p_1 + 3p_4 + 3p_7 \end{bmatrix}.
 \end{aligned}$$

例えば、状態2 (ノーアウトランナー1塁) において得点するのは、三塁打して本盗に成功する、または、本塁打のときに限り2点、二塁打して三盗失敗、二塁打して三盗せず、二塁打して三盗成功する、三塁打して本盗失敗、または、三塁打して本盗しない、のときに限り1点の得点を得るから、状態2における期待得点は $R(2) = 2(p_9 + p_{10}) + p_4 + p_5 + p_6 + p_7 + p_8$ となる。

3. 例

'96年の日本プロ野球セ・パ両リーグの打撃上位30傑の計算例を紹介する。盗塁に関しては、三盗、本盗、続けて2つの塁を奪うというデータが60人全員について揃わなかったため、盗塁は全て二盗とみなした。三盗、本盗、続けて2つの塁を奪うという割合は極めて低いので、TOERA値への影響はさほどないと思われる。Table.1, Table.2はセ・パ両リーグについて、入力データとOERA値とTOERA値の計算結果である。

Table 1. '96 セ・リーグ打撃30傑入力データ, OERA・TOERA 値

	氏名	AB	H	2B	3B	HR	BB	SB	LB	BA	OERA	TOERA
1	パウエル	518	176	42	2	14	46	1	3	.340	7.9255 (8)	7.7954 (8)
2	辻	400	133	9	2	2	53	9	3	.333	6.6587 (16)	6.6944 (17)
3	立浪	511	165	39	2	10	71	2	5	.323	7.8120 (9)	7.6185 (9)
4	山崎	453	146	20	0	39	46	1	1	.322	8.8030 (4)	8.7564 (3)
5	オマリー	461	145	23	0	18	75	3	4	.315	8.1007 (7)	7.9470 (7)
6	江藤	388	122	19	1	32	80	8	3	.314	10.4584 (1)	10.3399 (1)
7	松井	487	153	34	1	38	75	7	2	.314	9.5420 (2)	9.4960 (2)
8	西山	411	129	18	2	3	51	4	2	.314	6.1779 (21)	6.1595 (21)
9	前田	396	124	20	2	19	36	0	0	.313	7.1275 (10)	7.1275 (10)
10	ロペス	503	157	29	1	25	38	2	1	.312	6.9464 (13)	6.9271 (13)
11	稲葉	436	135	26	3	11	46	9	5	.310	6.6227 (17)	6.5280 (19)
12	ローズ	483	147	21	6	16	57	1	0	.304	6.7459 (15)	6.7586 (15)
13	コールズ	513	155	15	1	29	60	0	3	.302	7.0810 (12)	6.9587 (11)
14	落合	376	113	18	0	21	70	3	0	.301	8.4880 (5)	8.5280 (5)
15	金本	423	127	18	2	27	81	18	7	.300	8.8383 (3)	8.6723 (4)
16	谷繁	380	114	25	3	8	59	2	3	.300	7.0827 (11)	6.9584 (12)
17	駒田	485	145	22	1	10	46	1	3	.299	5.5747 (22)	5.4818 (22)
18	鈴木	355	106	15	0	13	50	6	2	.299	6.8273 (14)	6.8223 (14)
19	和田	520	155	22	3	5	43	2	7	.298	4.9574 (27)	4.7720 (28)
20	大豊	462	136	19	1	38	69	3	1	.294	8.3923 (6)	8.3717 (6)
21	マック	484	142	28	0	22	47	12	1	.293	6.4503 (19)	6.5520 (18)
22	野村	514	150	30	3	12	43	8	7	.292	5.5150 (23)	5.3816 (24)
23	飯田	424	123	19	3	6	24	13	9	.290	4.5338 (29)	4.4437 (29)
24	佐伯	390	113	23	5	6	35	3	4	.290	5.3744 (25)	5.2519 (25)
25	石井	496	140	19	3	1	82	45	21	.282	5.4013 (24)	5.4282 (23)
26	ブラッグス	356	100	20	1	13	50	6	1	.281	6.4108 (20)	6.4495 (20)
27	緒方	516	144	25	6	23	71	50	10	.279	6.5908 (18)	6.7526 (16)
28	土橋	370	103	18	3	6	36	7	1	.278	4.9119 (28)	4.9909 (27)
29	久慈	504	140	13	2	0	58	14	11	.278	4.1010 (30)	4.0432 (30)
30	中村	376	102	11	1	12	42	2	0	.271	5.0720 (26)	5.1054 (26)

AB: 打数, H: 安打数, 2B: 二塁打, 3B: 三塁打, HR: 本塁打, BB: 四死球, SB: 盗塁, LB: 盗塁死
 BA: 打率, OERA: OERA 値, TOERA: TOERA 値, () 内はその評価での順位

Table 2. '96 パ・リーグ打撃30傑入力データ, OERA・TOERA 値

	氏名	AB	H	2B	3B	HR	BB	SB	LB	BA	OERA	TOERA
1	イチロー	542	193	24	4	16	65	35	3	.356	8.8821 (1)	9.2045 (1)
2	片岡	416	131	20	3	15	66	3	3	.315	8.0200 (4)	7.9032 (4)
3	堀	465	145	28	3	16	56	8	2	.312	7.2964 (6)	7.3069 (6)
4	鈴木健	354	107	17	1	21	49	1	1	.302	7.8502 (5)	7.8020 (5)
5	秋山	466	140	27	0	9	41	13	4	.300	5.6404 (14)	5.6800 (14)
6	吉永	390	115	21	2	20	71	1	3	.295	8.1762 (3)	8.0221 (3)
7	ローズ	501	147	29	1	27	58	11	8	.293	7.0679 (7)	6.8613 (7)
8	村松	406	119	14	9	0	50	58	26	.293	5.2295 (16)	4.4839 (20)
9	松井	473	134	22	5	1	17	50	9	.283	3.6943 (27)	4.2754 (23)
10	水口	442	124	18	2	8	37	5	0	.281	4.6663 (20)	4.7432 (19)
11	田口	509	142	24	1	7	33	10	4	.279	4.2566 (22)	4.2847 (22)
12	田中	513	142	29	3	22	56	3	2	.277	6.0463 (13)	6.0069 (13)
13	ニール	430	118	24	0	32	82	1	2	.274	8.2294 (2)	8.1383 (2)
14	中村紀	411	112	15	1	26	42	4	1	.273	6.1739 (10)	6.1803 (10)
15	鈴木	426	114	21	1	9	52	3	4	.268	5.0287 (17)	4.9330 (17)
16	ジャクソン	489	130	21	3	19	32	10	4	.266	4.7765 (19)	4.7740 (18)
17	初芝	469	124	23	5	17	38	2	1	.264	4.9985 (18)	4.9884 (16)
18	南洵	439	116	22	1	5	48	7	1	.264	4.4045 (21)	4.4730 (21)
19	金子誠	395	103	14	2	4	28	15	2	.261	3.5587 (29)	3.7643 (27)
20	清原	487	125	30	0	31	76	0	2	.257	6.7589 (9)	6.6831 (9)
21	広瀬	390	100	21	3	3	34	8	2	.256	3.8884 (25)	3.9506 (25)
22	大島	397	101	11	3	4	55	8	2	.254	4.1961 (23)	4.2627 (24)
23	井出	414	105	13	5	7	36	11	10	.254	3.8667 (26)	3.7499 (28)
24	浜名	485	123	16	8	3	40	33	12	.254	3.5151 (30)	3.7027 (29)
25	垣内	387	98	19	3	28	40	16	3	.253	6.1256 (11)	6.1557 (11)
26	ブリトー	470	119	17	0	29	69	1	1	.253	6.1238 (12)	6.0975 (12)
27	平井	357	89	16	2	5	37	4	6	.249	3.9638 (24)	3.8300 (26)
28	小久保	478	118	26	3	24	55	7	3	.247	5.4259 (15)	5.3911 (15)
29	デューシー	427	105	17	5	26	88	3	3	.246	6.9368 (8)	6.8404 (8)
30	佐々木	374	91	11	0	9	30	7	4	.243	3.5673 (28)	3.5661 (30)

AB: 打数, H: 安打数, 2B: 二塁打, 3B: 三塁打, HR: 本塁打, BB: 四死球, SB: 盗塁, LB: 盗塁死
 BA: 打率, OERA: OERA 値, TOERA: TOERA 値, () 内はその評価での順位

幾つかの興味深い結果が見てとれる。

1. 金本・大豊・ニール・清原・デューシーらは打率順位が低いにもかかわらず、TOERA 値・OERA 値の順位が高い。彼らは本塁打が多く、得点への貢献が高い。
2. TOERA 値で9点以上は、江藤(10.3399), イチロー(9.2045), 松井(巨人)(9.4960)である。印象はイチロー・松井に比べ弱い、江藤の貢献度は著しい。
3. OERA 値によると松井(巨人)はイチローを0.6599ポイント上回るが、TOERA 値では0.2915まで差は縮まる。
4. OERA 値と TOERA 値による変動を見ると、イチロー(35盗塁,3盗塁死), 松井(西武)(50盗塁,9盗塁死)の値の変動が高く、それぞれ得点力において、約0.9ポイント、約0.6ポイント上昇する。

4. むすび

打者評価の方法としての TOERA 値を基に、年棒/TOERA 値による打者年棒評価、オールスターベストオーダーの選出、トレードに伴うチーム全体の TOERA 値の変動によるトレード効果分析等へ適用すると興味深いと思える。

謝辞

OERA 値、TOERA 値の計算を実行していただいた学部学生の裏川雅之、牛田悟郎、大坪雄一郎氏に感謝致します。

参考文献

- [1] Cover, T.M., and Keilers, C.W., An Offensive Earned-Run for Baseball, Operations Research, Vol.25(1977), 729-740.
- [2] ベースボール・マガジン社, 1996 ベースボール・レコードブック, ベースボール・マガジン社, 東京, 1997.
- [3] 木下栄蔵, マネジメントサイエンス入門, 啓学出版, 東京, 1993.
- [4] 竹内啓, 藤野和建, スポーツの数理科学, 共立出版, 東京, 1988.
- [5] Taylor, H.M., and Karlin, S., An Introduction to Stochastic Modeling, Revised Edition, Academic Press, San Diego, 1994.