reduced significantly for both the AG and the DG implementations by the precomputation of all sinusoidal terms that are required in (17), (18), and (20); the sinusoidal terms are computed before the call to LPNLP, and the resulting arrays of values are made available to subroutines FXNS and GRAD by the use of labeled common blocks.

Although the absolute maximum of the digital filter problem occurs at two points, there appeared to be no significant advantage for the solution to converge to one rather than the other. Thus, in Table X, the 1061 result was for Point A, and the 1055 result was for Point B.

V. Conclusion

Both constrained and unconstrained problems can be solved efficiently using gradient approximations. The problems considered are much easier to implement using the discrete gradient approach. This occurs because the often lengthy and error-prone process of deriving expressions for the exact gradient is bypassed. Also, for complicated gradient expressions, significant numerical errors can accumulate in the evaluation of the analytical gradient, as is evidenced by the results of the digital filter problem. Still, this enthusiasm for the discrete gradient approach must be tempered by the fact that CPU costs for many AG results can be somewhat lower than costs for corresponding DG results.

Birta's conclusion [3] that DFP was never improved by a periodic reset was based on insufficient data. For the unconstrained test problems considered, a significant improvement in efficiency was obtained by resetting the search direction to the gradient direction after every $2n + 1$ line searches. For the digital filter constrained problem, direction resets produced a definite stabilizing effect on the results. In contrast to DFP, SS was always hindered by periodic resets.

Convergence of LPNLP is based on gradient information. When using a gradient approximation, inaccuracies can cause problems with final convergence. This occurred for some test problems presented here. At an optimum, the gradient of the augmented Lagrangian should be zero. However, if an attempt is made to force the gradient components to be unrealistically small at the optimum, the algorithm may terminate with an error message, even though the optimum has been found for most practical purposes.

References

[1] R. Fletcher and M. J. D. Powell, "A rapidly convergent descent method for minimization," *Computer Journal*, vol. 7, pp. 163–168, July 1963.
[2] S. S. Oren, "On the selection of parameters in self-scaling variable metric algorithms," *Mathematical Programming*, vol. 7, pp. 351–367, Dec. 1974.
[3] L. G. Birta, "Some investigations in function minimization," *IEEE Trans. Systems, Man, Cybern.*, vol. SMC-6, pp. 186–197, Mar. 1976.
[4] D. A. Pierre and M. J. Lowe, *Mathematical Programming via Augmented Lagrangians: An Introduction with Computer Programs.* Reading, MA: Addison-Wesley, 1975.
[5] A. R. Colville, "A comparative study on nonlinear programming codes," IBM New York Scientific Center Report no. 320-2949, June 1968.
[6] M. J. D. Powell, "An iterative method for finding stationary values of a function of several variables," *Computer Journal*, vol. 5, pp. 147–151, July 1962.
[7] S. S. Oren, "Self-scaling variable metric (SSVM) algorithms, Part II: Implementation and experiments," *Management Science*, vol. 20, pp. 863–874, Jan. 1974.
[8] D. A. Pierre, "Multiplier algorithms for nonlinear programming," Eighth International Symposium on Mathematical Programming, Aug. 1973.
[9] J. B. Rosen and S. Suzuki, "Construction of nonlinear programming test problems," *Communications of the Association for Computing Machinery*, vol. 8, p. 113, Feb. 1965.
[10] M. J. D. Powell, "A method for nonlinear constraints in minimization problems, "in *Optimization*, R. Fletcher, Ed. New York: Academic Press, 1969.

Football and Basketball Predictions Using Least Squares

R. T. STEFANI

***Abstract*—Least squares is used to obtain ratings for college and professional football teams and for college basketball teams. Predictions are then made from the ratings. Observed accuracy was 72 percent for 3000 college football games, 68 percent for 1000 pro football games, and 69 percent for 2000 college basketball games.**

Introduction

Virtually everyone has attempted to predict the outcome of some sporting event. It should be no surprize that automated football ratings predate both high-speed digital computers and wire service ratings. In the 1930's, the P. B. Williamson "System" [1] was widely published. Williamson used such factors as hardness of schedule, gameness, and a "guts" factor combined by calculus and least squares. An intricate filing system implemented by a clerical staff (mini computers?) allowed him to rate 400 college teams. Dr. Litkenhaus of Vanderbilt University has provided computer-aided high school football and basketball ratings since 1941 for the Nashville Banner, resulting in a state championship trophy.

The Richard Dunkel "Index" was widely published from about 1955 until Mr. Dunkel's death in 1975. Dunkel updated his weekly ratings depending on the margin by which the higher rated team won or lost, with a higher penalty when loss occurred.

In 1969, Ed Mintz developed Compusport. Football predictions and ratings are carried by more than 25 newspapers. The factors were generated by trial and error using several hundred past game results. The predicted win margin is half the difference in team ratings plus a 2–4 point factor for home team advantage, advantage being greater for higher rated teams. Ratings are adjusted by comparing predicted point spread with actual point spread, and applying a threshold and a multiplying factor. Adjustment for loss may be greater than for victory. Variation is limited to ± 11 points. The Dunkel and Compusport adjustment is independent of the number of games played. A strong point in both is the relatively small adjustment when one-sided games occur or when a team performs way off of its normal rating. Compusport claims 75 percent accuracy in college football, 71–72 percent accuracy in pro football, and 55–58 percent correctness against the point spread. A similar performance-based algorithm was used by the International Chess Federation to establish all-time ratings for grand masters dating back to the 1700's.

The most sophisticated system is that of Bud Goode [2]. Goode has correlated past offensive and defensive statistics with subsequent victory. Those statistics with the highest correlation coefficients are blended to provide predictive capability. The system is not practical for small colleges where statistics are not readily available. A similar procedure was reported by Felson [3] for application to the stock market, where a multitude of statistics are available.

Little appears in the formal literature pertaining to rating procedures and resulting accuracy. This correspondence is unique in documenting a least squares system resulting in college and pro football predictions which have been published by the *Fort Worth Star-Telegram* since 1971. Basketball predictions

Manuscript received May 24, 1976.
The author is with the Department of Electrical Engineering, California State University, Long Beach, CA 90840.

have also been studied. This system requires only past scores and opponents and proceeds directly from classical least squares, no "learning" of algorithm factors being required. Accuracy is highly competitive with all predictions to which comparisons have been made.

Football Rating System

The model equation below relates the outcome of previously played games to team ratings subject to uncertainty (team performance varies randomly):

$$w = Ar + v \tag{1}$$

where

- w win margin column vector ($M \times 1$) of M games (w_k is the amount by which team i defeated team j in game k),
- A selection matrix ($M \times N$) (If team i defeated team j in game k, then $a_{k,i} = +1$ and $a_{k,j} = -1$. For ties $w_k = 0$, and either team may be designated the winner for entries into A.),
- r team rating column vector ($N \times 1$) where r_i refers to team i,
- v zero mean random noise vector ($M \times 1$) where v_k refers to the random error for game k.

In order to establish an estimate ($\hat{r}$) of the team rating vector, a least squares objective function is selected so that

$$J = e^T e \tag{2}$$

$$e = w - A\hat{r}. \tag{3}$$

The value of $\hat{r}$ which minimizes (2) must satisfy

$$A^T A\hat{r} = A^T w. \tag{4}$$

Before attempting to invert $A^T A$, inspection of the definition of A leads one to the conclusion that the

$$(i,j) \text{ element of } A^T A = \begin{cases} n(i), & i = j \\ -1, & i \neq j, \text{ but } i \text{ played } j \\ 0, & \text{otherwise} \end{cases}$$

where $n(i)$ = number of games played by team i. It follows that the sum of the elements of every row of $A^T A$ is zero, hence $A^T A$ is singular and noninvertible. One additional requirement could be made: the average rating must be 100. This would add one more row to A, a row of all ones; $w_{M+1} = 100N$, and A would be nonsingular. However, to analyze 120 college teams, a 120×120 matrix would have to be inverted. An equivalent, but computationally more desirable approach is taken.

The ith equation of the left-hand side of (4) represents $n(i)$ times the rating of team i minus the sum of team i's opponent's ratings. The right-hand side of the ith equation is team i's total win margin. If the rating of team i is solved for, one concludes that each team rating should be the average opponent rating plus the average win margin. This may be written compactly for all teams as

$$\hat{r} = D\hat{r} + \bar{w} \tag{5}$$

where

$$D = G^{-1}(G - A^T A),$$

= schedule matrix [(i,j) element is $+1/[n(i)]$ if team i played team j],

$$\bar{w} = G^{-1}A^T w = \text{average win margin vector},$$

$$G = \begin{bmatrix} n(1) & & & 0 \\ & n(2) & & \\ & & \ddots & \\ 0 & & & n(N) \end{bmatrix} = \text{games played matrix}.$$

Finally (5) may be solved iteratively

$$\hat{r}^{l+1} = D\hat{r}^l + \bar{w} \tag{6}$$

where $\hat{r}^l$ = lth estimate of the vector r. If all elements of r are 100 initially ($\hat{r}^0$), then (6), if it converges, results in the same answer as solving (4) with A augmented to require an average rating of 100. In any algorithm such as (6), important considerations are convergence (can a final value be approached) and acceleration (if the algorithm converges, can the number of iterations be reduced).

When a team has played few games, usually one or two, then (6) tends to oscillate. To eliminate this oscillation, the estimate $\hat{r}^1$ following the initial estimate $\hat{r}^0$ is adjusted by

$$\hat{r}^{1^1} = \tfrac{1}{2}(\hat{r}^1 + \hat{r}^0). \tag{7}$$

This adjustment only occurs for $\hat{r}^1$. The computational sequence is thus $\hat{r}^0$, $\hat{r}^1$, $\hat{r}^{1^1}$, $\hat{r}^2$, $\hat{r}^3$, etc. In six seasons using the adjustment of (7), (6) has always converged.

Convergence is accelerated by a wise choice of $\hat{r}^0$. It can be shown for round-robin play that the following initial rating for each team is immediately convergent (all subsequent estimates are the same), while the average rating is 100:

$$\hat{r}_i^0 = 100 + \bar{w}_i + \frac{1}{n(i)+1} \sum_{\substack{j=1 \\ j=i\text{'s} \\ \text{opponent}}}^{N} \bar{w}_j. \tag{8}$$

In words: a team's rating is 100 plus the team's average win margin plus $1/[n(i)+1]$ times the sum of each opponent's average win margin. If this were to be done for all teams,

$$\hat{r}^0 = 100 + \bar{w} + (G + I)^{-1}(G - A^T A)\bar{w}. \tag{9}$$

The computer is instructed to use (8) as the first estimate for teams who have played more than four games, while the previous season's rating is used otherwise. Use of (8) coupled with (7) results in an $\hat{r}^6$ that differs from $\hat{r}^5$ by less than 0.3 for nearly all teams. The sixth iteration is then used for prediction.

Data Processing

At the beginning of each season, schedules are encoded in character format, checked by a special program, and converted to numerical format for the matrix D in (6). Previous season's ratings are encoded for use as $\hat{r}_i^0$ in the early season. Weekly results of 120 college teams and 26 pro teams are reduced to win margins (w_k) and punched on eight data cards. The computer cross-checks the w_k, then updates $\bar{w}$ for the next week's run. Ratings are computed, and winners for the following weeks games are predicted by taking the rating differences. The program is written in Fortran IV for use on a CDC 3150 computer. Compilation and execution time is less than 1 min per weekly prediction. Accuracy is defined as the fraction by which the predicted winner actually wins. Ties are counted as $\frac{1}{2}$ right and $\frac{1}{2}$ wrong. The difference (in magnitude) between predicted point spread and actual point spread is noted and averaged.

Football Example

Suppose Arizona is designated team 1, Notre Dame is designated team 2, and USC is designated team 3. In three successive games; Notre Dame defeated USC by 14, Arizona defeated

USC by 10 and Notre Dame defeated Arizona by 8. The resulting matrices are, for $M = 3$ and $N = 3$,

$$A = \begin{bmatrix} 0 & 1 & -1 \\ 1 & 0 & -1 \\ -1 & 1 & 0 \end{bmatrix} \qquad w = \begin{bmatrix} 14 \\ 10 \\ 8 \end{bmatrix}$$

$$A^T A = \begin{bmatrix} 2 & -1 & -1 \\ -1 & 2 & -1 \\ -1 & -1 & 2 \end{bmatrix} \qquad G = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{bmatrix}$$

$$D = \begin{bmatrix} 0 & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & 0 & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} & 0 \end{bmatrix} \qquad \bar{w} = \begin{bmatrix} 1 \\ 11 \\ -12 \end{bmatrix}.$$

If (8) is used to define the initial rating for all three teams, it follows that (9), the vector form of (8), is applicable. As a result

$$\hat{r}^0 = \begin{bmatrix} 100\frac{2}{3} \\ 107\frac{1}{3} \\ 92 \end{bmatrix}.$$

Subsequent use of (6) results in a sequence of rating vectors equal to $\hat{r}^0$; thus convergence is immediate. Instead, if all elements of $\hat{r}^0$ are equal to 100, using (6), the following represents all but $\hat{r}^2$, $\hat{r}^3$, and $\hat{r}^4$:

$$\hat{r}^1 = \begin{bmatrix} 101 \\ 111 \\ 88 \end{bmatrix} \qquad \hat{r}^{1'} = \begin{bmatrix} 100.5 \\ 105.5 \\ 94 \end{bmatrix}$$

$$\hat{r}^5 = \begin{bmatrix} 100.66 \\ 107.22 \\ 92.12 \end{bmatrix} \qquad \hat{r}^6 = \begin{bmatrix} 100.67 \\ 107.39 \\ 91.93 \end{bmatrix}.$$

The fifth and sixth iterations are close enough together to select the sixth iteration for purposes of football estimation. Further, the sixth estimate is quite close to the extact rating vector above. If a rematch should occur between Notre Dame and USC, Notre Dame would be predicted to win by 15 points.

Football Results

Tables I and II show the results for 120 college football teams and all 26 pro football teams, respectively, over six seasons. In 3156 college football games, accuracy was 72 percent, while for 1043 pro games, the accuracy was 67.7 percent. Variation in accuracy from year to year is predictable from random chance. What is not predictable is that college and pro accuracy rise and fall together; in fact, the correlation coefficient is 0.8. The conclusion may be that pro and college football accuracy do not influence each other, but that each is correlated with macroscopic sociological phenomenon which, in turn, influences the overall feeling of well being and consistency of performance.

The best week in college football occurred during the 1975–1976 season (45 out of 52 for 86.5 percent). For three separate weeks in pro football, the computer has had 12 of 13 (92.3 percent).

The relationship between predicted point spread (rating difference) and accuracy is shown in Tables III and IV. In college football, the greater the difference in ratings between opponents, the greater is the accuracy. In pro football, however, accuracy viewed versus predicted point spread peaks and then drops off for higher predicted point spreads. This anomaly is probably due to lack of talent in college football where, spirit aside, low-rated teams only cause major upsets 10 percent of the time. In pro football, a team experiencing a poor season has the talent, but not teamwork. Against a highly rated team, the talent pulls together to cause major upsets 30 percent of the time. Note also that the number of college games versus predicted point spread is relatively uniform except for a drop-off at high predicted point spreads. In pro football, there are progressively fewer games at higher point spreads.

The home team advantage was studied for games picked by the computer to be within five points. In 300 college games, the home team won 61 percent, while in 150 pro games, the home team won 56 percent. The computer, ignoring home team advantage, was accurate in 54 percent of the college games and 57 percent of the pro games. Considerable improvement in college accuracy would occur if a home team advantage is considered, but pro accuracy would actually be reduced somewhat.

During the 1970–1971 season, 10 of the 120 college football teams were small west-coast colleges. During the 1971–1972 and 1972–1973 seasons, 10 of the teams were small Texas colleges (Lone Star Conference). In each case, the accuracy was not significantly different than that of major colleges. The method is applicable to all levels of football: major, minor, community college, and high school.

Comparison as to accuracy has been made against human selectors (wire service, sports writers) and computer selectors, (Dunkel, Goode, Compusport). Nearly 2500 comparisons have been made. Least squares maintained a five game lead overall. No significant difference in accuracy was observed against individual human selectors, against any of the computer selectors (some using more sophisticated strategies), or against an overall combination thereof.

In gambling, where legal, of course, football cards are played (for example, "Notre Dame over USC by 7"). The gambler may select Notre Dame and be correct if Notre Dame wins by 8 or more. Conversely, he may take USC and be correct if either USC wins or if USC loses by 6 or less. The gambler is always wrong if Notre Dame wins by 7. To win money, the gambler must select a number of games on a card and be correct 4 out of 4, 5 out of 5, etc. The payoff increases with the number of games he chooses as in Table V. However, note the required point spread selection correctness to break even. Selecting four games requires the least correctness, hence the greatest expected return. The least squares predictions have been compared with 600 alternative point spreads. The computer correctly gave, or took, the points 55.3 percent of the time, but, on cards in which four selections are made, 56.2 percent correctness is required to break even. Thus more computational power than that of least squares (or predictors that have been compared to least squares) is needed to beat football cards even under the most favorable circumstances. The computer nearly breaks even which answers the gambler's prayer, "Lord, let me break even, I need the money."

College Basketball Rating System

College basketball teams play about 25 games per season, many opponents being unknown at the season's start because of extensive tournament play. Uncertainty in scheduling makes (6) inappropriate, since both D and $\bar{w}$ would have to be updated weekly. Instead of adjusting the entire $\hat{r}$ vector as games are played, only the ratings for opponents in each game are adjusted, other team ratings being held constant. In order to maintain the

TABLE I
COLLEGE FOOTBALL ACCURACY PER SEASON

Season	Games	Right	Wrong	Accuracy	Average Error Magnitude of Predicted Point Spread	Champion	Rating
1970–1971	232	171	61	0.738	13.6	Texas	131.2
1971–1972	581	409	172	0.705	13.4	Nebraska	136.7
1972–1973	560	409½	150½	0.732	14.0	Nebraska	135.6
1973–1974	595	438½	156½	0.737	14.7	Oklahoma	134.8
1974–1975	598	411	187	0.687	15.9	Oklahoma	139.4
1975–1976	590	434	156	0.736	13.7	Oklahoma	125.8
Total	3156	2273	883	0.720	14.3		

TABLE II
PRO FOOTBALL ACCURACY PER SEASON

Season	Games	Right	Wrong	Accuracy	Average Error Magnitude of Predicted Point Spread	Champion	Rating
1970–1971	98	68½	29½	0.700	12.2	Minnesota	114.1
1971–1972	189	120½	68½	0.638	12.4	Dallas	109.9
1972–1973	189	128½	60½	0.680	12.2	Miami	111.2
1973–1974	189	130½	58½	0.690	13.2	Miami	114.8
1974–1975	189	124½	64½	0.659	11.2	Washington	109.0
1975–1976	189	134	55	0.709	12.3	Pittsburgh	114.6
Total	1043	706½	336½	0.677	12.3		

TABLE III
COLLEGE FOOTBALL ACCURACY PER PREDICTED POINT SPREAD

Predicted Point Spread	Games	Right	Wrong	Accuracy
0–5	663	378½	284½	0.571
6–11	649	396	253	0.610
12–13	683	492	191	0.720
19–23	635	538½	96½	0.848
29 and up	526	468	58	0.890

TABLE IV
PRO FOOTBALL ACCURACY PER PREDICTED POINT SPREAD

Predicted Point Spread	Games	Right	Wrong	Accuracy
0–5	323	194½	128½	0.602
6–10	276	183	93	0.663
11–16	241	184	57	0.763
17–26	136	98½	37½	0.724
27 and up	67	46½	20½	0.694

TABLE V
FOOTBALL CARDS

Selections	Payoff for \$1 Bet if 100% Correct	Correctness per Game to Break Even
3	5	0.585
4	10	0.562
5	15	0.581
6	25	0.585
7	35	0.604
8	50	0.613
9	75	0.618
10	100	0.631

least squares nature of (6), ratings for team i and team j, opponents in game $M + 1$, must satisfy

$$\hat{r}_i = \frac{1}{n(i) + 1}\left[\hat{r}_j + w_{M+1} + n(i)\hat{r}_{i\,\text{old}}\right] \tag{10}$$

$$\hat{r}_j = \frac{1}{n(j) + 1}\left[\hat{r}_i - w_{M+1} + n(j)\hat{r}_{j\,\text{old}}\right] \tag{11}$$

where $\hat{r}_{i\,\text{old}}$ and $\hat{r}_{j\,\text{old}}$ satisfied (6) for $n(i)$ and $n(j)$ games, respectively. w_{M+1} = amount by which team i defeated team j in game $M + 1$.

Simultaneous solution of (10) and (11) yields new ratings $\hat{r}_i$ and $\hat{r}_j$ where $n(i)$ and $n(j)$ are assumed equal:

$$\hat{r}_i = \hat{r}_{i\,\text{old}} + \frac{w_{M+1} - (\hat{r}_{i\,\text{old}} - \hat{r}_{j\,\text{old}})}{n(i) + 2} \tag{12}$$

$$\hat{r}_j = \hat{r}_{j\,\text{old}} - \frac{w_{M+1} - (\hat{r}_{i\,\text{old}} - \hat{r}_{j\,\text{old}})}{n(j) + 2}. \tag{13}$$

Equations (12) and (13) are used even when $n(i)$ and $n(j)$ are unequal. In a study of half of a pro football season, (12) and (13) yielded essentially the same accuracy as (6).

COLLEGE BASKETBALL EXAMPLE

Suppose Arizona, alphabetically the second team, with a rating of 116 won by 24 over UCLA, alphabetically the 112th team, with a rating of 102. Suppose both teams had played eight games prior to their encounter. Their new ratings would be, respectively, $\hat{r}_2 = 116 + 1 = 117$ and $\hat{r}_{112} = 102 - 1 = 101$. In a rematch, Arizona would be a 16 point favorite.

COLLEGE BASKETBALL RESULTS

Table VI shows college basketball results for 130 teams over two seasons. In 1926 games, the computer exhibited 69 percent accuracy. The average error in predicting the point spread was

TABLE VI
COLLEGE BASKETBALL ACCURACY PER SEASON

Season	Games	Right	Wrong	Accuracy	Average Error Magnitude of Predicted Point Spread	Champion	Rating
1972–1973	876	596	280	0.680	9.5	UCLA	124.5
1973–1974	1050	734	316	0.699	10.0	UCLA	121.9
Total	1926	1330	596	0.690	9.8		

9.5 points per game. Because of the large number of games per season, predictions seldom appear in print, except for tournament games at the season's end. Statistically significant comparisons therefore were not possible.

CONCLUSIONS

It was shown that the least square procedure, implemented on a digital computer, can achieve accurate predictions in college football, pro football, and college basketball.

ACKNOWLEDGMENT

The author acknowledges the encouragement and support of the *Fort Worth Star-Telegram* readers and, especially, of Sports Editors Herb Owens and Bob Lindley.

REFERENCES

[1] W. McG. Keefe, "Viewing the news," *New Orleans Times Picayune*, Section 4, pp. 2 and 4, Dec. 23, 1934.
[2] C. Gewecke, "The computer likes the Cowboys," *The Christian Science Monitor*, p. 10, Jan. 13, 1971.
[3] J. Felson, "Learning pattern recognition techniques applied to stock market forecasting," *IEEE Trans. Systems, Man, Cybern.*, vol. SMC-5, no. 6, pp. 583–594, Nov. 1975.

Book Reviews

Prozessidentifikation—Identifikation und Parameterschätzung dynamischer Prozesse mit diskreten Signalen (Process Identification—Identification and Parameter Estimation of Dynamic Processes with Discrete Signals)—R. Isermann (Berlin: Springer-Verlag, 1974, 188 pp.). *Reviewed by George M. Siouris, Aerospace Guidance and Metrology Center, Newark Air Force Station, Newark, OH 43055.*

Since the first IFAC Symposium on identification held in Prague, Czechoslovakia, in 1967, the field of identification and parameter estimation has come of age. This is attested by the great number of papers that have been published in recent years.

The present book is applications oriented. Emphasis is placed on small digital computer applications and the determination of the mathematical model for the static and dynamic behavior from input–output measurements. The identification and parameter estimation methods treated can be evaluated on a small digital computer and are suitable for a number of applications, including biological and economic processes. Plant modeling and identification usually require a great deal of effort before the convenient linear approximations are known. The selection of these plant models depend to a large degree on the available *a priori* information about the plant. If this information is not available, recourse is made to experimental techniques on the model. The model may be derived from basic principles, *a priori* knowledge, or statistical analysis of data. With the model comes an understanding of the process, the relative effect of various process parameters, and their interactions. There are 16 chapters in all.

Chapter 1 introduces the subject, states the problems of process identification, and classifies the identification procedures. Chapters 2 and 3 treat time-discrete signals and processes, and correlation analysis. Treated here are the usual topics of deterministic and stochastic signals and processes, stationary processes, autocorrelation and cross-correlation functions, correlation analysis of linear dynamic processes, and binary test signals. These two chapters may be omitted by readers familiar with the content matter, while they may serve as a good review for other readers. Chapter 4 covers the least squares method. Topics addressed include static and dynamic processes, recursive least squares, and weighted least squares. Chapter 5 is devoted to stochastic approximation. Unlike the recursive least squares, the stochastic approximation is more amenable to numerical computation. With this method one can obtain recursive algorithms when the on-line estimate is difficult to derive. The next chapter, Chapter 6, titled "Generalized Least Squares Method," is a natural extension of Chapter 4, and discusses briefly one-stage and recursive generalized least squares.

In Chapter 7 the reader will find a short description on instrumental variables. As in the previous chapter, the material is divided into one-stage and recursive cases. Chapters 8 and 9 are devoted to the maximum likelihood estimation and the Bayesian estimation theory. The various parameter estimation methods and their relationship with one another discussed up to this point are summarized in Chapter 10. Chapters 11–13 treat the least squares as applied to response functions in deterministic test signals, stochastic approximation, and correlation analysis, respectively. Chapter 14 compares the parameter estimation methods. After an overview of the various parameter estimation methods, the remainder of the chapter is devoted to the comparison of performance indices and gives numerical examples of selected processes. The chapter concludes with the graphical representation of six parameter estimation methods for three error model processes. Much of the material in this chapter is based on the author's work presented at the Third IFAC-Symposium on Identification. Chapter 15 contains useful information of the model order and dead time. Two numerical illustrative examples are given, and the results are shown in graphical form. The last chapter, Chapter 16, treats various problems such as choice of input signals, choice of the sampling time, the