Residuals for relative risk regression

BY WILLIAM E. BARLOW
Department of Preventive Medicine, University of Southern California, Los Angeles, California 90033, U.S.A.

AND ROSS L. PRENTICE
Fred Hutchinson Cancer Research Center, Seattle, Washington 98104, U.S.A.

SUMMARY
Several possible definitions of residuals are given for relative risk regression with time-varying covariates. Each such residual has a representation as an estimator of a stochastic integral with respect to the martingale arising from a subject’s failure time counting process. Previously proposed residuals for individual study subjects and for specific time points are shown to be special cases of this definition, as are previously derived regression diagnostics. An illustration and various generalizations are also given.

Some key words: Case-control study; Censoring; Failure time data; Influence function; Proportional hazards; Regression diagnostic; Relative risk regression; Residual.

1. Introduction
Relative risk regression models, based on the model of Cox (1972), have become extremely popular for the analysis of failure time data; see Prentice & Farewell (1986) for a recent review. As with standard linear regression models convenient and understandable methods are required to detect various departures from modelling assumptions. Suitably defined residuals may play an important role in such identification. However, the nonparametric aspect of the model, the possibility that modelled regression variables may be varying with follow-up time and, most importantly, the usual presence of right censorship, implies that specialized residual definitions are required. The class of residuals considered here is most easily formulated using counting process notation for the failure time data.

Let $N_i(t)$, $Y_i(t)$ and $Z_i(t)$ represent, respectively, for the $i$th subject the values of counting, censoring and covariate processes at follow-up time $t$ $(i = 1, \ldots, n)$, while \{ $N_i(u)$, $Y_i(u)$, $Z_i(u)$; $0 \leq u < t$ \} specifies the corresponding histories for the $i$th subject prior to time $t$. Thus in a typical univariate failure time application, $N_i$, with right-continuous sample paths, will take value zero prior to the time of failure on the $i$th subject and value one thereafter, while $Y_i$ with left-continuous sample paths will take value one at times at which the $i$th subject is ‘at risk’ for an observed failure, and value zero otherwise. The counting process $N_i$ can be uniquely decomposed so that for all $(t, i)$

$$N_i(t) = \Lambda_i(t) + M_i(t),$$

with cumulative intensity process $\Lambda_i(t) = \int_0^t \lambda_i(u) \, du$, where the integral is over $0, t$), and where $M_i$ is a local square integrable martingale (Andersen & Gill, 1982). Under standard independent failure time and independent censorship assumptions (Self & Prentice, 1982)
the intensity process, \( \lambda_t \), is the familiar ‘hazard’ function. A relative risk regression model (Cox, 1972) for this intensity process can then be written

\[
\lambda_i(t) = Y_i(t)\lambda_0(t)r\{z_i(t)\beta\},
\]

where \( z_i(t) \) is a row \( p \)-vector of functions of the \( i \)th subject’s preceding covariate history, or products of such functions and time, \( \beta \) is a corresponding column \( p \)-vector of parameters to be estimated, \( r \) is a fixed function usually taken to be \( r(v) = e^v \), \( r(v) = 1 + v \), or a mixture of these two functions (Thomas, 1981; Guerrero & Johnson, 1982; Breslow & Storer, 1985), and \( \lambda_0 \) is a fixed function giving the hazard function for a subject with a standard covariate history for which \( z(t) = 0 \). The sample paths of \( z_i \) are required to be left continuous with right-hand limits, so that \( z_i \) is ‘predictable’.

The log partial likelihood function (Cox, 1972, 1975) for estimating \( \beta \) can be written in stochastic integral notation as

\[
\int_0^{t_0} \sum_{i=1}^n \left\{ \log r_i(t) - \log \sum_{i=1}^n Y_i(t)r_i(t) \right\} dN_i(t),
\]

where \( r_i(t) = r\{z_i(t)\beta\} \) for \( i = 1, \ldots, n \) and \( t_0 \) is the maximum follow-up time for the sample. Differentiating with respect to \( \beta \) gives the ‘score statistic’ that can be rewritten as

\[
\int_0^{t_0} \sum_{i=1}^n x_i(t)\{dN_i(t) - p_i(t) d\tilde{N}(t)\},
\]

where \( x_i(t) = z_i(t)u^{(1)}\{z_i(t)\beta\} \), with \( u^{(1)}(x) = \partial \log r(x)/\partial x \), can be thought of as the ‘functional’ covariate for the \( i \)th subject at time \( t \),

\[
p_i(t) = Y_i(t)r_i(t) / \left\{ \sum_{i=1}^n Y_i(t)r_i(t) \right\}
\]

and \( \tilde{N} = N_1 + \ldots + N_n \). Andersen & Gill (1982) and Prentice & Self (1983) show the maximum partial likelihood estimate \( \hat{\beta} \), defined as a \( \beta \)-value such that (4) equals zero, to be consistent and asymptotically normal under mild conditions. Furthermore, \( \hat{\beta} \) is asymptotically jointly Gaussian with an estimator \( \hat{\Lambda}_0 \) of the cumulative baseline hazard function given by

\[
\hat{\Lambda}_0(t) = \int_0^t \left\{ \sum_{i=1}^n Y_i(u)\hat{r}_i(u) \right\}^{-1} d\tilde{N}(u),
\]

where \( \hat{r}_i(u) = r\{z_i(u)\hat{\beta}\} \).

2. RESIDUALS FOR RELATIVE RISK REGRESSION

Models of the form (2) specify a relative risk process, \( r\{z(t)\beta\} \), at each follow-up time \( t \), and are nonparametric in other respects. Accordingly, residuals should be directed toward the identification of greater than expected departures between the data and some aspect of the relative risk model. In particular, such departures may reflect inadequate modelling of the dependence of the relative risk on certain covariates; inadequate modelling of the relative risk in certain time periods; unsuitable choice of relative risk form \( r(v) \); or inadequate correspondence between the specified model and the accumulated data on single individuals or sets of individuals. Residuals may even play a role in an attempt to assess globally goodness-of-fit. For residuals to be useful one requires at
least approximate knowledge of their distribution under the specified modelling assumptions. For practical purposes it is often sufficient that residuals be centred about zero with known scaling.

As a potential residual for the ith subject, motivated by (1), consider

$$e_i(f) = \int_0^t f_i(t) \, dN_i(t) - \int_0^t f_i(t) \, d\Lambda_i(t) = \int_0^t f_i(t) \, dM_i(t), \quad (6)$$

for a predictable process $f_i$ with $f_i(t)$ defined in terms of data on the ith subject, and possibly other, subjects prior to time $t$. This random variable measures the discrepancy between the stochastic integral of $f_i$ with respect to $N_i$ and a corresponding integrated conditional expectation under the assumed model. In view of (6) this quantity is a martingale and hence has mean zero. Furthermore, since the predictable covariation process for $M_i$ and $M_j$ is zero ($i \neq j$) it follows that $e_i(f)$ and $e_j(f)$ are uncorrelated (Gill, 1980). The predictable covariation process corresponding to $e_i(f)$ is

$$\int_0^t f_i(t) \otimes^2 \lambda_i(t) \, dt, \quad (7)$$

where $a \otimes^2$ indicates the outer product of a vector $a$, with typical $(i, j)$th element $a_ia_j$.

The estimated residual corresponding to (6) can be written

$$\hat{e}_i(f) = \int_0^t \hat{f}_i(t) \, dN_i(t) - \int_0^t \hat{f}_i(t) \, d\hat{\Lambda}_i(t) = \int_0^t \hat{f}_i(t)\{dN_i(t) - \hat{\rho}_i(t) \, d\hat{N}(t)\} \quad (8)$$

using (5), where

$$\hat{\rho}_i(t) = Y_i(t)\hat{\tau}_i(t)/\sum_{i=1}^n Y_i(t)\hat{\tau}_i(t).$$

Furthermore, in view of (7), a variance estimator for (6) can be written

$$\hat{V}(\hat{f}) = n^{-1} \sum_{i=1}^n \int_0^t \hat{f}_i(t) \otimes^2 d\hat{\Lambda}_i(t) = n^{-1} \int_0^t \sum_{i=1}^n \hat{f}_i(t) \otimes^2 \hat{\rho}_i(t) \, d\hat{N}(t). \quad (9)$$

In this circumstance simple martingale convergence results show that asymptotically $\hat{e}_i(\hat{f})$ has mean zero, has variance estimated by (9), and is uncorrelated with $e_j(\hat{f})$, for $j \neq i$.

Consider now some special cases of (8) and their relationship to previously defined residuals and diagnostics. The special case

$$\hat{e}_i(1) = N_i(t_0) - \hat{\Lambda}_i(t_0) = \delta_i - \sum_{i=1}^n \delta_i \hat{\rho}_i(t_i)$$

contrasts the observed number of failures $N_i(t_0)$ on the ith subject with the sum of the conditionally expected number of failures over the follow-up period, where $t_i$ is the time of failure ($\delta_i = 1$) or terminal right censorship ($\delta_i = 0$) for the ith subject. In the special case of fixed covariates, $z_i(t) = z$ and simple right censorship, one can write $\Lambda_i(t_0) = \Lambda_0(t_i)r(z, \beta)$ for $i = 1, \ldots, n$. These quantities constitute a censored sample from a unit
exponential distribution. On the basis of this known distribution several authors (Crowley & Hu, 1977; Kay, 1977; Kalbfleisch & Prentice, 1980, pp. 96–8) have proposed \( \hat{\Lambda}_i(t_0) \) as a, possibly censored, residual for the \( i \)th subject. The above residual \( \hat{\varepsilon}_i(1) \) merely adds the conditional expectation of unity to the censored \( \hat{\Lambda}_i(t) \) values, changes sign and centres the resulting quantities about zero. A residual that compares the number of failures on the \( i \)th subject with a corresponding anticipated value under the model may not be sufficiently sensitive for the purposes mentioned at the beginning of this section. Accordingly, B. E. Storer, in a University of Washington Ph.D. dissertation, has shown by simulation that deliberate misspecification of a model, for example fitting an unimportant covariate while ignoring an important one, may yield a set of censored data residuals \( \hat{\Lambda}_i(t_0) \) from which the estimated survivor function differs less from exponentiality than does that derived under the correct model.

As a second special case consider

\[
\hat{\varepsilon}_i(t) = t_i \delta_i - \sum_{l=1}^{n} t_l \delta_l \hat{\beta}_l(t_l).
\]

This is a more traditional residual giving the difference between the failure time \( t_i \) for an uncensored subject and a corresponding projected quantity under the model. Individuals with small failure times will tend to have positive residuals, while those with large failure times will typically yield negative residuals. The residuals for censored individuals cannot be positive and are strictly less than those for other subjects having the same covariate history and an earlier uncensored failure time.

Since the models in question specify a relative risk, rather than a failure time model, a more natural definition may involve some comparison of realized and model projected relative risks for the \( i \)th subject. Hence, one may consider a residual

\[
\hat{\varepsilon}_i \{ \log \hat{r}_i(t) \} = \log \{ \hat{r}_i(t) \} \delta_i - \sum_{l=1}^{n} \log \{ \hat{r}_i(t_l) \} \delta_l \hat{\beta}_l(t_l),
\]

where logarithms have been taken to avoid range restrictions. This residual compares the magnitude of the log relative risk at the time of failure for the \( i \)th subject to a corresponding model derived value. This use of the relative risk size should increase its sensitivity, compared to \( \hat{\varepsilon}_i(1) \) for example. In contrast to \( \hat{\varepsilon}_i(t) \) the formulation in terms of relative risk preserves a functional invariance under monotone increasing transformations in failure time that seems fundamental to the model. Specifically a transformation from \( t \) to \( t' \), while requiring \( z'(t') = z(t) \) as with fixed covariates, will leave \( \hat{\varepsilon}_i \{ \log \hat{r}_i(t) \} \) unchanged.

For the standard exponential relative risk form, \( \log r_i(t) = z_i(t) \beta \), one may prefer to instead define a residual \( p \)-vector \( \hat{\varepsilon}_i \{ z_i(t) \} \) for the \( i \)th subject.

More generally, one could define a \( p \)-vector in terms of the functional covariate \( x_i(t) \) as

\[
\hat{\varepsilon}_i \{ \hat{x}_i(t) \} = \int_0^{t_0} \hat{x}_i(t) \{ dN_i(t) - \hat{\beta}_i(t) \hat{d}N_i(t) \} = \hat{x}_i(t_0) \delta_i - \sum_{l=1}^{n} \hat{x}_i(t_l) \delta_l \hat{\beta}_l(t_l).
\]  

(10)

As the derivative with respect to \( \hat{\beta} \) of the previous residual this quantity focuses on the influence or leverage that the \( i \)th subject has on the relative risk parameter estimate. A comparison of (10) with (4) shows that (10) can be thought of as the contribution that the \( i \)th subject makes to the score statistic evaluated at \( \hat{\beta} \). As such, it is not surprising
that it approximates, aside from standardization previously derived, the empirical influence functions. In fact, a refined residual definition that contrasts the difference between the functional covariate and its risk set mean with a corresponding projected difference; that is,

\[ \hat{e}_i \{ \hat{x}_i(t) - \hat{E}(\hat{\beta}, t) \} = \{ \hat{x}_i(t_i) - \hat{E}(\hat{\beta}, t_i) \} \delta_i - \sum_{i=1}^{n} \{ x_i(t_i) - \hat{E}(\hat{\beta}, t_i) \} \delta_i \hat{p}_i(t_i) \]  

is, aside from standardization, precisely the empirical influence of Cain & Lange (1984) and Reid & Crépeau (1985). Therefore (11) generalized their results to time-dependent covariates and to a general relative risk form. This residual also relates closely to the deletion diagnostic of Storer & Crowley (1985).

The fact that (11) focuses on differences between \( x_i(t) \) and the corresponding risk set mean

\[ \hat{E}(\hat{\beta}, t) = \sum_{i=1}^{n} Y_i(t) \hat{x}_i(t) \hat{p}_i(t) \]

would seem to add to its purity relative to (10), as a measure of discrepancy from model projections. One wonders, for example, whether the interpretation of (10) could be confounded by rapidly changing risk set means across time. Furthermore, the variance estimator (9) corresponding to (11) is simply \( n^{-1} \) times the score statistic variance variance matrix

\[ V(\hat{\beta}) = \int_0^{t_0} \sum_{i=1}^{n} [x_i(t) - \hat{E}(\hat{\beta}, t)] \hat{\delta}_i \hat{p}_i(t) d\hat{N}(t), \]

which is routinely computed in model fitting.

Corresponding to any of the above definitions residuals can be defined for any subset of the follow-up interval merely by restricting the range of the integration in (8) to the subset and by correspondingly restricting the time interval in the variance formula (9). One can even, at least formally, define a residual for each subject at each observed failure time. Corresponding to (8) such a residual for the \( i \)th subject at failure time \( t \) \( \{ d\hat{N}(t) \neq 0 \} \) is merely

\[ \hat{f}_i(t) \{ d\hat{N}_i(t) - \hat{p}_i(t) d\hat{N}(t) \}, \]  

with corresponding variance 'estimator'

\[ V(\hat{\beta}, t) = n^{-1} \sum_{i=1}^{n} \hat{f}_i(t) \hat{\delta}_i \hat{p}_i(t) d\hat{N}(t). \]

The residuals (12) may be summed over failure times to give (8). They may also be summed over study subjects at a fixed failure time to assess the agreement of the data at time \( t \) with the fitted relative risk model. The special case

\[ \{ \hat{x}_i(t) - \hat{E}(\hat{\beta}, t) \} \{ d\hat{N}_i(t) - \hat{p}_i(t) d\hat{N}(t) \} \]

is worth setting out specifically. Summing these quantities over study subjects gives the contribution at time \( t \) to the score statistic (4) evaluated at \( \hat{\beta} \). This generalizes the 'partial residuals' of Schoenfeld (1982) to time-dependent covariates and to relative risk forms other than the exponential. Note that these generalized Schoenfeld residuals pertain to time points rather than to individuals.
3. Further results

The above notation is general enough to allow tied failure times. Of course \( \hat{\beta} \) and \( \hat{\lambda}_0 \) and hence the above estimated residuals can be expected to encompass increasing asymptotic bias as the frequency of ties increases.

An important generalization of (2) allows stratification of the baseline hazard function \( \lambda_0 \). Residuals corresponding to (8) or (12) can be defined separately for each stratum. Different residual patterns among strata may reflect the need to allow the regression parameter \( \beta \) to vary among strata.

The counting process notation and partial likelihood estimation procedures are also general enough to allow multivariate failure times, assuming the intensity model (2) is applicable beyond the first failure time on a subject to second and subsequent failures. Residuals of the form (8) or (12) and their corresponding asymptotic variance estimators are still applicable. The merits of specific residual definitions, that is of the choice of \( f_i \), may differ from the above univariate case. For example, \( \hat{e}_i(1) \) which contrasts the number of failures on subject \( i \) with a corresponding model projected number may be a quite useful residual if individual study subjects tend to experience a large number of failures. With samples involving a large number of subjects each with a modest number of failure times a generalization of (2) to allow the subject to move to stratum \( s+1 \) immediately after experiencing an \( s \)th failure may often be natural (Prentice, Williams & Peterson, 1981), in which case univariate residuals of the type discussed in § 2 may be considered for each stratum.

The computation of (8) or (12) is a simple matter upon storing the distinct failure times \( t_1, \ldots, t_k \) in the sample, along with the failure multiplicity \( m_i \) and the relative risk total \( \sum Y_i(t_i) \hat{\beta}(t_i) \), where the sum is over \( i=1, \ldots, k \). For the special case (11) or (14) it is also necessary to store \( \hat{E}(\beta, t_i) \) at each such failure time. The residual for the \( i \)th subject can then be calculated by referring only to data on the \( i \)th subject.

Residuals of the form (12) may also be considered in the context of time-matched case-control sampling. Specifically the \( i \)th subject in a matched set with case occurrence at time \( t \) is \( \hat{f}_i(t)\{y_i - \hat{\beta}_i(t)\} \), where \( y_i = 1 \) for the case and zero for the corresponding matched controls. The special case

\[
\left\{ \hat{e}_i(t) - \sum_{l=1}^n \hat{e}_i(t)\hat{\beta}_i(t) \right\}\{y_i - \hat{\beta}_i(t)\} \tag{15}
\]

where \( l=1, \ldots, n \) indexes the matched set, may be particularly pertinent. Aside from standardization (15) with \( y_i = 0 \) approximates the effect of the deletion of a control subject on \( \hat{\beta} \), while the sum of (15) over subjects in the matched set approximates the effect of case deletion on \( \hat{\beta} \), since the matched set ceases to contribute to the likelihood function following case deletion. Similar expressions are given by Moolgavkar, Lustbader & Venzon (1984). Lustbader (1986) proposes another residual for this context. Aside from standardization it is given simply by \( y_i - \hat{\beta}_i(t) \).

Vector residuals such as (11) may be standardized by dividing the \( j \)th component by the square root of the \((j, j)\)th element of the corresponding variance estimator, \( V(\hat{\beta})_{jj} \). Such a variance estimator will typically be nonzero, though not with (12) if the \( j \)th components of \( f_i(t) \) are identical for all subjects at risk at \( t \). Then the \( j \)th component of (12) is zero at that time point, so that the corresponding standardized value can also be taken to be zero.
In some circumstances it may be desirable to define a scalar ‘global’ residual corresponding to a vector residual definition. One possibility, corresponding to (11),

\[ \int_0^{t_0} \{ \hat{x}(t) - \hat{E}(\hat{\beta}, t) \} V(\hat{\beta})^{-1} \{ \hat{x}(t) - \hat{E}(\hat{\beta}, t) \}^T d\hat{M}_i(t) \]  

(16)
is included in the following illustration.

Plots of such scalar residuals versus a suitably selected covariate may provide insight into the choice of model form. For example with fixed covariates, Barlow (1985) indicates that the adequacy of an exponential form model relative to a linear model with the same regression variables may be examined by adding the additional covariate \( \log (1 + z_i\hat{\beta}) \) or \( (1 + z_i\hat{\beta}) \log (1 + z_i\hat{\beta}) \) to the model and testing for a zero value of the corresponding parameter. Alternatively one may plot a global residual, such as (16), versus the values of the additional covariate in order to detect a trend.

4. ILLUSTRATION

Consider an example in which a standard exponential form relative risk model may not apply. Specifically consider the subsample of the lung cancer data (Kalbfleisch & Prentice, 1980, Appen. 1) given by Reid & Crépeau (1985). There are 37 failure times, two of which are tied, among the 40 observations. Suppose only two covariates are included in the model, performance, PERF, status and an indicator of adeno, ADENO, tumour cell type. Both covariates are significant in a standard exponential form relative risk model with fixed covariates. However, fitting the time-dependent product term \( \text{ADENO} \times t \) yields a likelihood ratio test value of 4.03 with 1 degree of freedom. Therefore, the proportional hazards assumption may not hold. Hopefully, violation of this assumption will be indicated by the residual pattern when both covariates are included in the model.

The residuals described below are restricted to the special case \( f_i(t) = \hat{x}_i(t) - \hat{E}(\hat{\beta}, t) \) with \( \hat{x}_i(t) = z_i \). Figure 1(a) shows the unstandardized residual components (14) for the covariate ADENO plotted against the rank of failure or censoring time. The triangles indicate such components for subjects failing at time \( t \) while the plus signs indicate components for the corresponding nonfailures. Since all subjects with adeno tumour type fail within 90 days there is no variability in ADENO after this time. There appears to be a time trend in the residuals since most of the early residuals are slightly negative. If the residual components are summed over individuals at each failure time the Schoenfeld residual is obtained. These residuals are plotted against time in Fig. 1(b). Summing over time, instead of individuals, yields the unstandardized integrated residuals (11) shown in Fig. 1(c). Since the individuals are ordered by failure or censoring time a relation of ADENO with time seems clear and the residuals appear to indicate the proportional hazards model may not hold. A plot of the empirical influence function is expected to show a similar pattern since the covariance of the two estimated coefficients is low. The scalar measure of influence of Cain & Lange (1984) is shown in Fig. 1(d). This plot shows the large influence of one of the censored observations which has a large impact on the estimate associated with PERF. The global residuals (16) were also computed and are plotted against time in Fig. 1(e). These residuals show a decreasing pattern in time. Figure 1(f) plots these global residuals against \( \log (1 + z^2) \) with \( z_1 = 90 - \text{PERF} \) and \( z_2 = \text{ADENO} \). The transformation in performance status is necessary to guarantee the
Fig. 1. Lung cancer data. (a) Residual component versus rank of failure or censoring time; triangles, failures at that time. (b) Schoenfeld residual versus rank of failure time. (c) Unstandardized integrated residual versus rank of failure, shown by triangle, or censoring, shown by plus sign, time. (d) Scalar influence measure versus rank of failure, shown by a triangle, or censoring, shown by plus sign, time. (e) Estimated integrated global residual versus rank of failure, shown by triangle, or censoring, shown by plus sign, time. (f) Estimated integrated global residual versus log \((1 + z'\theta)\); triangles, individuals who failed; plus signs, censored individuals.

positivity of the argument and only the sign of the first coefficient is changed with the magnitudes unaffected. No particular pattern is observed suggesting the exponential form may be appropriate with these regression variables. In some cases a 3-dimensional plot of the residual component against time and the covariate might be useful, particularly if the covariate is time-dependent.
5. Discussion

A general discussion of possible residuals for relative risk regression has been given. Though it would be of interest to compare the merits of competing choices for the ‘discrepancy process’ \( \hat{f}_i \), the special case \( \hat{f}_i(t) = \hat{x}_i(t) - \hat{E}(\hat{\beta}, t) \) seems to have particular merit in view of the close connection to empirical influence and regression diagnostics. For this special case the proposed residuals can be viewed, aside from standardization, as a generalization to time-dependent covariates and general relative risk form of the empirical influences of Cain & Lange (1984) and Reid & Crépeau (1985). A generalization of Schoenfeld’s (1982) partial residual is obtained by summing the residuals at a particular time point over individuals. It seems preferable to plot the unstandardized integrated residual (11) rather than the corresponding empirical influences since the former is less affected by the covariance among the regression parameter estimates. For example, Reid & Crépeau (1985) find two observations have large influence on the coefficients of PERF and ADENO when all seven covariates are in the model. When only the two significant covariates are included observation 30 has negligible influence on the coefficient of ADENO. It appears that the apparent influence of observation 30 was induced by the covariance of the estimates of ADENO, SMALL and SQUAMOUS.

The types of residuals defined here can complement techniques based on the generalization and fitting of models of the form (2) in the process of building an appropriate relative risk model. Certainly there are many paths to a model that appears to fit. For example, a more complex choice of modelled regression variable can offset an unfortunate choice of relative risk form. Hence the goal of the use of residuals, and of model fitting more generally, includes parsimony and interpretability in addition to adequacy of fit.

Acknowledgement

The authors would like to thank a referee for several helpful comments.

References


[Received August 1986. Revised May 1987]