Numerical Weather Prediction

The effects of nonlinearity on analysis and retrieval errors

Forecasting Research Technical Report No. 252

Jonathan Eyre

email: nwp_publications@metoffice.com

Forecasting Research
Technical Report No. 252

THE EFFECTS OF NONLINEARITY ON ANALYSIS AND RETRIEVAL ERRORS

by

J R Eyre

October 1998

Meteorological Office
NWP Division
Room 344
London Road
Bracknell
Berkshire
RG12 2SZ
United Kingdom

Tel: 44 (0)1344 856245 Fax: 44 (0)1344 854026 e-mail: jsarmstrong@meto.gov.uk

THE EFFECTS OF NONLINEARITY ON ANALYSIS AND RETRIEVAL ERRORS

by J R Eyre

Abstract

The error characteristics of an optimal retrieval/analysis are well understood for the linear problem, and linear theory has also been widely used to assess the approximate error characteristics of weakly nonlinear problems. However, in the context of preparations for advanced infra-red sounders, it has become apparent that linear theory will lead to an over-optimistic assessment of performance, and particularly of the potential contribution of water vapour channels to the accuracy of temperature retrieval/analysis.

In this paper, the effects of nonlinearities in the radiative transfer on retrieval/analysis errors are investigated theoretically for the general variational retrieval/assimilation problem. It is shown that these effects can be treated as an additional source of "forward model" error, which we call nonlinearity error. The distribution of nonlinearity error is inherently non-Gaussian and is a potential source of significant bias in the retrieval/analysis. Also, the magnitudes of analysis error and nonlinearity error are mutually dependent. Strategies for quantifying this error source are proposed. They are relevant not only to the assessment of the information content of advanced sounder data but also to the mitigation of nonlinear effects during retrieval or assimilation.

1. Introduction

When analyzing meteorological fields, it is important to consider sources of significant error, both in the observations available to the analysis and in the analysis procedure itself. For mathematically optimal results, it is necessary to account correctly for the error characteristics of the observations and of any other information used in the analysis. For practical applications, where sub-optimal results are usually acceptable, it is nevertheless important to characterize significant error sources approximately; serious deficiencies in the analysis usually result if an important source of error is overlooked or greatly under-estimated.

Amato et al. (1996) have drawn attention to a particular problem caused by nonlinear relationships between some observed quantities and those to be analyzed. They have demonstrated detrimental effects in the case of retrieving atmospheric temperature profiles from spectra of infra-red radiances. Similar effects are potential sources of significant error for general problems of this type: when retrieving/analyzing a set of geophysical variables from a set of observations related to them in a nonlinear way (see Joiner and da Silva 1998; Stoffelen 1998, section VI.1.2).

This paper examines the errors introduced by these nonlinearities within analysis/retrieval schemes that are formulated to be optimal in the absence of such effects. In particular, it explores the theoretical error characteristics of nonlinear variational methods now in use for analysis of the 3D atmospheric state from global observations (e.g. Andersson et al. 1994) or the retrieval of atmospheric temperature and humidity profiles from satellite sounding radiances (e.g. Eyre et al. 1993). It discusses how these errors will tend to manifest themselves and suggests some approaches to estimating their effects, and to mitigating them.

2. Theory

2.1 Solution of the variational problem

The problem of variational analysis is usually posed as one of minimising a penalty function containing a number of quadratic terms, each representing a source of information on the variables to be analyzed (e.g. see Lorenc 1986, Le Dimet and Talagrand 1986). Using the notation of Ide et al. (1997), the penalty function for a combination of "observed" information with "background" information is:

$$J(x) = \tfrac{1}{2}(x-x^b)^T.B^{-1}.(x-x^b) + \tfrac{1}{2}(y^o-H\{x\})^T.(E+F)^{-1}.(y^o-H\{x\}) \tag{1}$$

where x a vector representing the state to be analyzed,
x^b is an estimate of x obtained from "prior" or "background" information,
B is the error covariance of x^b,
y^o is a vector of observations,
$H\{x\}$ is a vector containing equivalent values in the observation space corresponding to the state x and is calculated through a "forward model" (or "observation operator"), $H\{...\}$,
E is the error covariance of the observations, and
F is the error covariance of the forward model.
^{T} and $^{-1}$ denote matrix transpose and inverse respectively.

In general, the retrieval/analysis problem may be nonlinear in several ways. For example, B, E or F could be a function of the state x or the measurements y^o. However, the only potential nonlinearity represented in eq.(1) is in the observation operator $H\{x\}$, and only this aspect of nonlinearity is considered in this paper.

Eq.(1) is minimised by solving its gradient equation:

$$B^{-1}.(x-x^b) - H'(x)^T.(E+F)^{-1}.(y^o-H\{x\}) = 0, \tag{2}$$

where $H'(x) = \nabla_x H\{x\}$. We define the solution of this equation as the analysis, $x=x^a$. Then

$${}^{a}-x^b) = H'(x^a)^T.(E+F)^{-1}.(y^o-H\{x^a\})). \tag{3}$$

Lorenc (1986) shows that this solution is the most probable solution for x (i.e. the mode of its probability density function), if the error characteristics of observations and background are Gaussian, and observation errors are uncorrelated with background errors. There is no assumption of a linear observation operator in the derivation of the maximum probability solution; $H\{x\}$ can be a nonlinear function of x. However, this solution of the problem will only be truly optimal if the error characteristics of the real problem correspond to those we assume in the analysis of the data.

Using ^{t} to denote true values, we can re-write eq.(3) as

$$B^{-1}.(x^a-x^t-x^b+x^t) = H'(x^a)^T.(E+F)^{-1}.(y-y^t-H\{x^a\}+y^t). \tag{4}$$

Introducing ε^a, ε^b and ε^y for the errors in x^a, x^b and y^o respectively, and defining forward model error, $\varepsilon^f = H\{x^t\}-y^t$,

$$B^{-1}.(\varepsilon^a-\varepsilon^b) = H'(x^a)^T.(E+F)^{-1}.(\varepsilon^y-\varepsilon^f-H\{x^a\}+H\{x^t\}). \tag{5}$$

2.2 The linear case

For the linear problem, i.e. where $H\{x\}$ is a linear function of x, and so $H'(x) = H' =$ a constant, we can solve eq.(3) analytically to give

$$x^a = x^b + K.(y^o-H\{x^b\}) \tag{6}$$

where $K = B.H'^T.(H'.B.H'^T+E+F)^{-1}$ (7)

x^a is "optimal" in the following respects: if x^b and y^o are unbiased, then x^a is unbiased and has minimum error variance; additionally, if the errors of x^b and y^o are Gaussian, x^a also corresponds to the maximum likelihood solution.

It can be shown (see Rodgers 1976) that the covariance of ε^a is given by

$$A^{-} = B^{-1} + H'^T.(E+F)^{-1}.H' \tag{8}$$

or

$$A = B - K.H'.B = (I-K.H').B\ . \tag{9}$$

It can also be shown (from eq.(5), or see Eyre 1987) that

$$\varepsilon^a = A.B^{-1}.\varepsilon^b + A.H'^T.(E+F)^{-1}.(\varepsilon^y-\varepsilon^f)$$

or

$$\varepsilon^a = (I-K.H').\varepsilon^b + K.(\varepsilon^y-\varepsilon^f).$$

From this equation for the error in a single retrieval/analysis, an equivalent equation for the mean error of an ensemble of retrievals/analyses can be obtained in terms of the means of the various components of error.

$$\langle\varepsilon^a\rangle = (I-K.H').\langle\varepsilon^b\rangle + K.(\langle\varepsilon^y\rangle-\langle\varepsilon^f\rangle),$$

where <...> denotes a mean value.

Clearly, if ε^b, ε^y and ε^f are all unbiased, then ε^a will also be unbiased; otherwise eq.(11a) gives us a method for assessing the propagation of biases from different sources into the solution.

Let us denote the mean error and the covariance of error for the linear case, given by equations (8) and (11a), as $<\varepsilon^a{}_L>$ and A_L respectively.

2.3 The case of the nonlinear observation operator

When y(x) is nonlinear, two types of problem arise. The first type concerns the practical problems of finding the solution x^a. Minimisation methods for solution of eq.(2) or (3), or their equivalents, tend to become slower to converge as nonlinearity increases. Also the nonlinearity may introduce multiple minima, making it more difficult to find the global minimum. These may, indeed, be real problems in practice but they are not the primary concern in this paper. We shall assume that x^a, corresponding to a suitable minimum of J(x), can be found and that we are concerned with the second type of problem, namely the error characteristics of the solution x^a.

With the introduction of nonlinearity, the modal solution of eq.(2) or (3) is no longer the minimum variance solution. However, it is not our intention here to re-examine the various possible definitions of the "optimal" solution. We accept that the solution of eq.(2) is our "best" estimate, but we are interested to understand its error characteristics.

Following Amato et al. (1996) we expand the forward model *H*(x) as a Taylor series (except that here we do so around the analyzed value x^a):

$$H\{x^t\} = H\{x^a\} + H'(x^a).(x^t-x^a) + \text{higher order terms in } (x^t-x^a)$$

The higher order terms give rise to a "nonlinearity error" ε^n which we define by:

$$H\{x^a\} - H\{x^t\} = H'(x^a).(x^a-x^t) + \varepsilon^n = H'(x^a).\varepsilon^a + \varepsilon^n.$$

ε^n arises because $H'(x^a)$ differs from $H'(x^t)$. Figure 1 illustrates a scalar example.

Substituting eq.(13) into eq.(5) gives

$$B^{-1}.(\varepsilon^a-\varepsilon^b) = H'(x^a)^T.(E+F)^{-1}.(\varepsilon^y-\varepsilon^f-H'(x^a).\varepsilon^a-\varepsilon^n).$$

or

$$(B^{-1}+H'(x^a)^T.(E+F)^{-1}.H'(x^a)).\varepsilon^a = B^{-1}.\varepsilon^b + H'(x^a)^T.(E+F)^{-1}.(\varepsilon^y-\varepsilon^f-\varepsilon^n),$$

or

$$\varepsilon^a = A_L(x^a).B^{-1}.\varepsilon^b + A_L(x^a).H'(x^a)^T.(E+F)^{-1}.(\varepsilon^y-\varepsilon^f-\varepsilon^n).$$

Note that eq.(16) is the analogue of eq.(10) for the linear case, but with the inclusion of ε^n in the last term, and so can be written:

$$<\varepsilon^a> \; = \; <\varepsilon^a{}_L> + A_L(x^a).H'(x^a)^T.(E+F)^{-1}.<\varepsilon^n>$$

At this point it is possible simply to absorb the nonlinearity error ε^n as a component of the forward model error ε^f. Then equations (8) to (11a) can be applied to the nonlinear case. However, because the characteristics of the nonlinearity error are different from those of other sources of forward model error (as will be shown below in section 3.1), it is instructive to keep ε^n separate.

The mean error and covariance of error for the nonlinear case can now be expressed in terms of their linear counterparts as follows:

$$<\varepsilon^a> \; = \; <\varepsilon^a{}_L> + K(x^a).<\varepsilon^n>$$

and, assuming that nonlinearity error is uncorrelated with other components of error,

$$A \; = \; A_L + K(x^a).N.K(x^a)^T$$

where N is the covariance of nonlinearity error. The validity of this assumption is discussed in section 3.1.

Equations (18) and (19) show how the nonlinearity error (in the measurement space) is mapped by the retrieval/analysis operator $K(x^a)$ into addition terms contributing to the mean error in the retrieval/analysis and its error covariance respectively. Note that presence of this new source of error does not (in general) change the expression representing the modal retrieval/analysis, eq.(3). However, it does change the value of the error in this analysis.

Alternatively, as suggested above, nonlinearity error may be treated as a new component of forward model error and simply absorbed into F. This is closer to "optimal", in the sense that the assumed errors now correspond more closely to the actual errors, but in doing so the question of "optimality" must be re-examined for the reasons discussed in section 3.2.

3. Discussion

3 Error distributions and the introduction of bias

Equations (12) and (13) show that the nonlinearity error ε^n represents quadratic and higher order terms in ε^a in the expansion of $H\{x^t\}$. If the quadratic term is dominant, then ε^n will be quadratic in ε^a. If all other error sources are Gaussian then, in the linear case, ε^a will also be Gaussian. Therefore, in the nonlinear case, ε^n will **not** have a Gaussian distribution. In fact, the quadratic relation between ε^a and ε^n will give the same sign to values of ε^n for all values of ε^a, and thus the mean value of ε^n will be non-zero.

The effect of non-zero bias in ε^n, if uncorrected, will cause a bias in the analysis/retrieval through eq.(18), even if there are no other sources of bias. (i.e. $<\varepsilon^a{}_L>=0$). Such biases have been observed in practice for the problem of retrieving wind speed and direction from scatterometer data (see Stoffelen and Anderson 1997). [For the scatterometer retrieval problem, there is an additional complication: although the forward model errors are close to

Gaussian in the retrieval space, they become highly non-Gaussian when mapped into the observation space, and so the solution of eq.(2) or (3) no longer represents the most probable solution.]

3.2 Error covariances and optimality assumptions

If the quadratic dependence of ε^n on ε^a is dominant, this will ensure that there is no first order correlation between ε^n and ε^a, and hence between ε^a and other sources of error, thus justifying the assumptions behind eq.(19), at least approximately.

Because the nonlinearity error will usually have a dominant quadratic term, the magnitude of the additional analysis error should vary as the square of the analysis error itself, at least when the effect of nonlinearity is relatively small.

Values of forward model error covariance that are appropriate for the linear case will not be appropriate to achieve a solution of minimum error variance for the nonlinear case. Even if the mean bias is corrected (e.g. empirically), the analysis will still have an error covariance given by eq.(19). This will not only represent error varainces higher than for the linear case A_L, but higher than if the forward model error covariance is increased to account for nonlinearity error (i.e. where the error covariance is given by eq.(8) in which F is replaced by F+N). If F is adjusted in this way and the mean bias corrected, then solution accuracy will be improved. However, because the analysis error and nonlinearity error are mutually dependent (through equations (13) and (16)), the covariance of the latter cannot now be estimated analytically (but see section 3.3).

Also, because the distribution of the nonlinearity error is inherently non-Gaussian, eq.(3) will no longer represent the exact maximum likelihood solution. To achieve this, we would have to construct a penalty function containing a term appropriate to the particular distribution of the nonlinearity error (i.e. proportional to the log of its probability density function - see, e.g., Lorenc 1986, Ingleby and Lorenc 1993, Andersson and Jarvinen 1999). Therefore, although eq.(3) may represent an adequate approximation to the maximum likelihood solution, it no longer represents a formally optimal solution.

3.3 Assessment and mitigation of nonlinear effects

The significance of nonlinearity error will vary greatly from problem to problem. It is essentially related to the change in the value of H' for typical values of analysis error. Note that there is a significant difference here from the problem presented by Amato et al. (1996): in their linearized one-step retrieval, the nonlinearity error is related to the background error, whereas in the fully nonlinear variational framework it is related to the analysis/retrieval error, which can be much smaller. (For example, for a retrieval using climatological background information, as illustrated by Amato et al., the retrieval error will be much smaller than the background error.)

When an observation is known to be related nonlinearly to the analysis variables, the magnitude of the nonlinearity error should be investigated. Significant degradation of performance may be expected when the nonlinearity error exceeds the combined measurement and forward model error but is not accounted for in the assumed statistics.

In practice, if forward model error covariance is deduced empirically (e.g. from observation-background statistics), it is likely to include contributions from nonlinearity error. Similarly, the biases derived from observation-background statistics will contain contributions from nonlinearity error, in addition to any caused by measurement, forward model or background biases. Empirical bias tuning procedures based on observation-background statistics will tend to compensate for these effects. However, they will not do so exactly, since the magnitudes of nonlinearity effects associated with background errors will be different from those associated with analysis errors. It is therefore important to consider in which space (analysis or measurement) and for which variable (background or analysis) it is most desirbale that the results be unbiased.

It will be particularly important to assess and allow for the effects of nonlinearity error when other sources of noise are low, and particularly (as shown by Amato et al.) when background constraints are non-existent or very weak.

The effects of nonlinearity error may be quantified approximately as follows:

(a) Evaluate analysis error covariance in the linear limit using eq.(8).

(b) Use this value to estimate nonlinearity error through eq.(13), i.e. to obtain an estimate for a typical case or an ensemble of cases.

(c) Add the estimate of nonlinearity error to F and re-evaluate the analysis error using eq.(8).

(d) If significantly different, repeat (b) and (c).

(e) Evaluate the mean analysis error with eq.(18) using a revised value for $K(x^a)$ consistent with the new value of F.

This procedure for iterative estimation of nonlinearity effects is illustrated in Fig.2. For the well-behaved functions shown here, it appears to be a convergent process. Fig.2 also suggests that the process could be initiated at (b), with nonlinearity error estimated from the background error (in place of analysis error in the linear limit).

4. Conclusions

The presence of nonlinearity in the relationship between observed and analyzed variables is a source of bias in the analysis and of increased error variance. The potential significance of these effects will need to be studied case by case, but significant effects are to be expected when the nonlinearity error is comparable to or greater than other sources of measurement and forward model error. Moreover the magnitude of the nonlinearity error is related to the analysis error itself. Section 3.3 suggests a recipe for assessing quantitatively the effects of nonlinearity error.

References

Amato, U., De Feis, I. and Serio C., 1996.
Linearization pseudo-noise and its effect on the retrieval of atmospheric state from infra-red spectral radiances.
Geophys. Res. Lett., **23**, 2565-2568.

Andersson, E. and Jarvinen, H., 1999.
Variational quality control.
To appear in Q. J. R. Meteorol. Soc. (Also ECMWF Tech Memo 250, 1998.)

Andersson, E., Pailleux, J., Thépaut, J.-N., Eyre, J. R., McNally, A. P., Kelly, G. A. and Courtier, P., 1994.
Use of cloud-cleared radiances in three/four dimensional variational data assimilation.
Q. J. R. Meteorol. Soc., **120**, 627-653.

Eyre, J. R., 1987.
On systematic errors in satellite sounding products and their climatological mean values.
Q. J. R. Meteorol. Soc., **113**, 279-292.

Eyre, J. R., Kelly, G. A., McNally, A. P., Andersson, A. and Persson, A., 1993.
Assimilation of TOVS radiance information through one-dimensional variational analysis.
Q. J. R. Meteorol. Soc., **119**, 1427-1463.

Ide, K., Courtier, P., Ghil, M. and Lorenc, A. C., 1997.
Unified notation for data assimilation: operational, sequential and variational
J. Meteorol. Soc. Japan, **75**, 181-189.

Ingleby, N. B. and Lorenc, A. C., 1993.
Bayesian quality control using multivariate normal distributions.
Q. J. R. Meteorol. Soc., **119**, 1195-1225.

Joiner, J. and da Silva, A. M., 1998.
Efficient methods to assimilate remotely sensed data based on information content.
Q. J. R. Meteorol. Soc., **124**, 1669-1694.

Le Dimet, F. X. and Talagrand, O., 1986.
Variational algorithms for analysis and assimilation of meteorological observations: theoretical aspects.
Tellus, **38A**, 97-100.

Lorenc, A. C., 1986.
Analysis methods for numerical weather prediction.
Q. J. R. Meteorol. Soc., **112**, 1177-1194.

Rodgers, C. D., 1976.
Retrieval of atmospheric temperature and composition from remote measurements of thermal radiation.

Geophys. Space Phys., **14**, 609-624.

Stoffelen, A., 1998.
Scatterometry.
Doctoral thesis, University of Utrecht (ISBN 90-393-1708-9).

Stoffelen, A. and Anderson D., 1997.
Ambiguity removal and assimilation of scatterometer data.
Q. J. R. Meteorol. Soc., **123**, 491-518.

Figures

1 Illustrating the source of nonlinearity error for a scalar case.

2. Illustrating, for a scalar case, the iterative estimation of nonlinearity error N and analysis error A, starting either from the analysis error in the linear limit A_L or from the background error B.

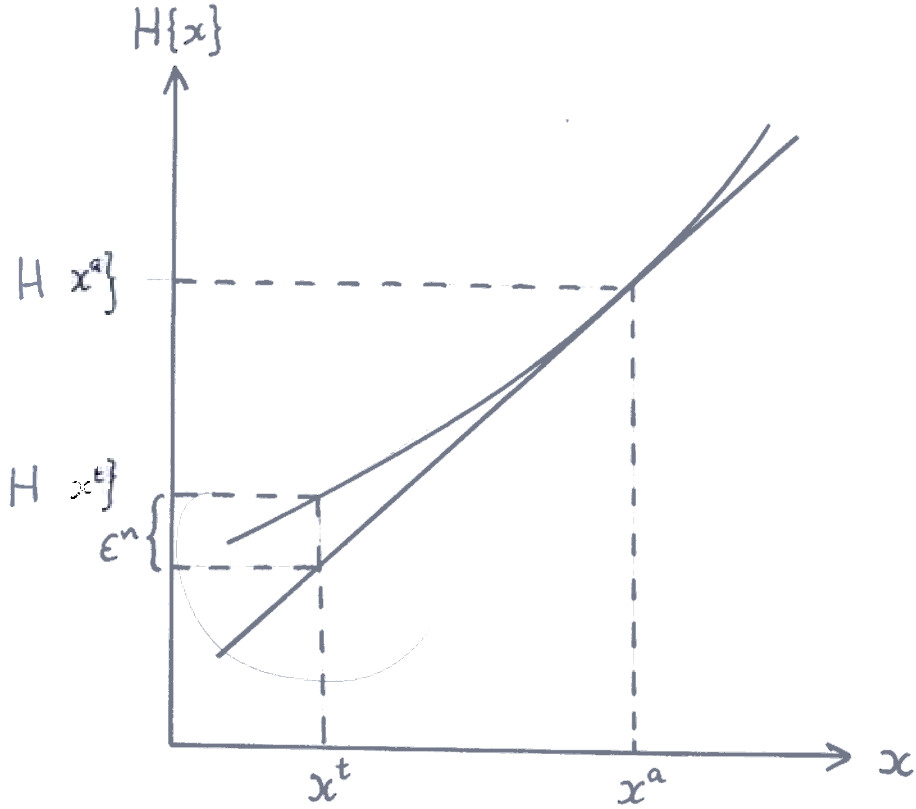
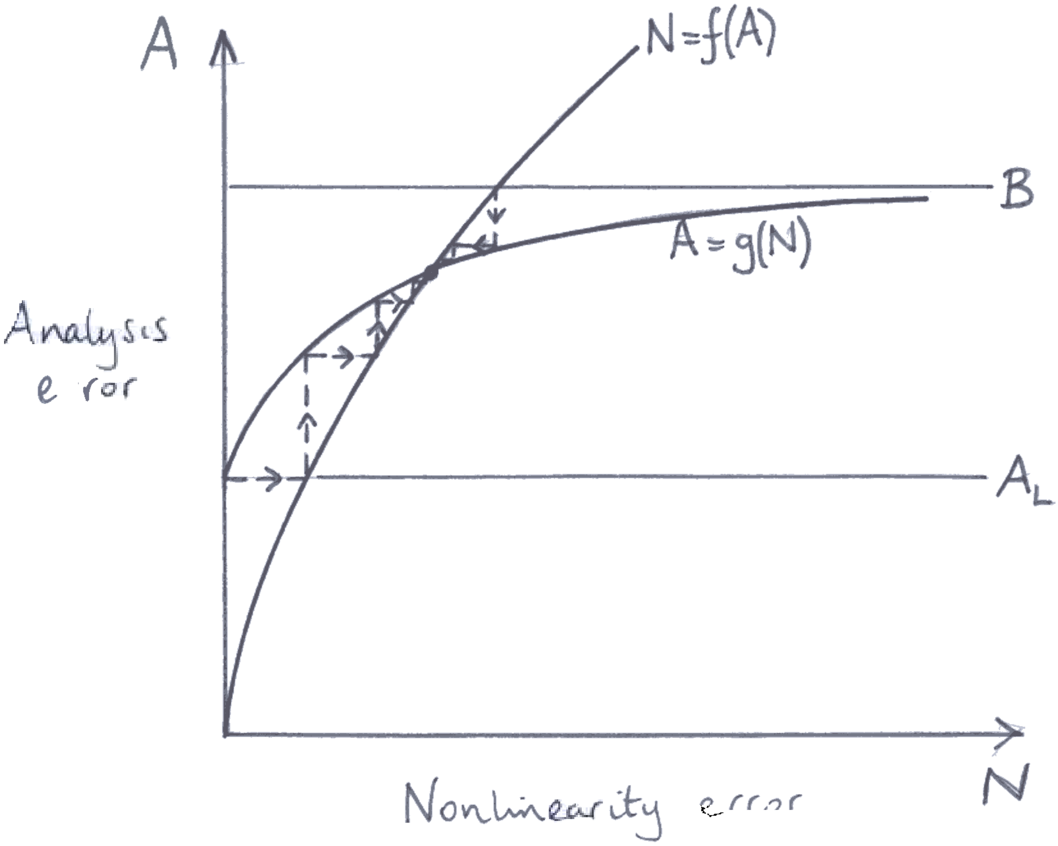


Figure 1 Illustrating the source of nonlinearity error for a scalar case



<u>Figure 2</u> Illustrating the iterative estimation of nonlinearity error N and analysis error A starting either from the analysis er. the linear limit A_L or from the background error B