Drawbacks of a Method for Calculationg Pseudoinverses

Söderström, Torsten

1971

Document Version:
Publisher's PDF, also known as Version of record

Link to publication

Citation for published version (APA):
Söderström, T. (1971). *Drawbacks of a Method for Calculationg Pseudoinverses*. (Research Reports TFRT-3035). Department of Automatic Control, Lund Institute of Technology (LTH).

Total number of authors:
1

LUND UNIVERSITY

PO Box 117
221 00 Lund
+46 46-222 00 00

DRAWBACKS OF A METHOD FOR CALCULATING PSEUDOINVERSES.

TORSTEN SÖDERSTRÖM

REPORT 7113(B) SEPTEMBER 1971
LUND INSTITUTE OF TECHNOLOGY
DIVISION OF AUTOMATIC CONTROL

DRAWBACKS OF A METHOD FOR CALCULATING PSEUDOINVERSES.

ABSTRACT - An iterative method for computing pseudoinverses given by Ben-Israel and Cohen is proved to be unstable if the rank of the given matrix is not maximal. The unstable case is illustrated by numerical examples.

This work was supported by the Swedish Board for Technical Development (contract 71-50/U33.

1. INTRODUCTION.

The pseudoinverse of a matrix may be defined in different (but equivalent) ways [6]. There are several algorithms for computing pseudoinverses. In [5] and [6] some comparisons between different methods are made. The purpose of this paper is to analyze a method given by Ben-Israel and Cohen [1].

2. THE ALGORITHM.

In [1] the following iterative algorithm is given. The pseudoinverse of A is denoted by $A^{\dagger}$.

$$Y_{k+1} = Y_k(2I-AY_k) \qquad k = 0, 1, \ldots \tag{1}$$

$$Y_0 = \alpha A^T \tag{2}$$

If

$$0 < \alpha < \frac{2}{\max \lambda(A^TA)} \tag{3}$$

it is shown that $Y_k \rightarrow A^{\dagger}$ as $k \rightarrow \infty$.

Further it is shown

$$|| A^{\dagger} - Y_{k+1} || \leq || A || \cdot || A^{\dagger} - Y_k ||^2 \tag{4}$$

In [1] there is also a corresponding method for computing $AA^{\dagger}$.

$$Z_{k+1} = 2Z_k - Z_k^2 \qquad k = 0, 1, \ldots \tag{5}$$

$$Z_0 = \alpha AA^T \tag{6}$$

If (3) is fulfilled $Z_k \to AA^{\dagger}$ as $k \to \infty$. Further tr Z_k is a monotone increasing sequence which converges to rank A.

3. SOME QUESTIONS ABOUT THE ALGORITHM.

Some practical questions concerning the method are not quite satisfactorily answered.

1. How difficult is it to determine the rank A correctly? It is well-known, [5], that the problem of determining the rank of a matrix may be difficult when uncertainties and/or roundoff errors have to be considered.
2. The method involves numerous matrix multiplications. What will this mean in terms of computing time and accuracy?
3. How critical is the choice of α? The optimal value of α is shown to be

   $$\alpha = 2\left[\max \lambda(A^TA) + \min_{\lambda \neq 0} \lambda(A^TA)\right]^{-1}$$

   but the computation of this quantity seems to require roughly as much work as the computation of the singular value decomposition [4]. This latter method is believed to be one of the best available. A comparison between this method and another is given in [6] by numerical examples.

It is believed that the numerical examples toward the end of this paper will shed some light upon these questions.

From (4) one could expect that (1) could be used for improvements of a solution. However, this is not true since the algorithm is unstable if rank A is not maximal.

4. A TRANSFORMATION OF THE EQUATIONS.

In order to demonstrate the unstability of the solutions of (1) and (5), introduce

$$X_k = A^\dagger - Y_k \tag{7}$$

$$T_k = AA^\dagger - Z_k \tag{8}$$

Equations (1) and (5) are then transformed to

$$X_{k+1} = A_1X_k + X_kA_2 + X_kAX_k \tag{9}$$

$$T_{k+1} = A_2T_k + T_kA_2 + T_k^2 \tag{10}$$

$$A_1 = I - A^\dagger A \tag{11}$$

$$A_2 = I - AA^\dagger \tag{12}$$

It is interesting to examine the stability properties of the solutions $X_k = 0$ of (9) and $T_k = 0$ of (10). The starting values corresponding to (2), (6) are

$$X_0 = A^\dagger - \alpha A^T \tag{13}$$

$$T_0 = AA^\dagger - \alpha AA^T \tag{14}$$

5. SOME LEMMAS.

In the following the singular valve decomposition will be used.

LEMMA 1. *Let A be a real m×n matrix. Then there is a factorization*

$$A = U\Sigma V^T \tag{15}$$

U m×m orthogonal matrix, V n×n orthogonal matrix, Σ m×n diagonal matrix. $\Sigma = \mathrm{diag}(\sigma_1 \ldots \sigma_p)$, $p = \min(m,n)$. The numbers σ_i are the singular values. Further the pseudoinverse is given by

$$A^\dagger = V\Sigma^\dagger U^T \tag{16}$$

where $\Sigma^\dagger$ is the n×m matrix $\mathrm{diag}(\sigma_1^\dagger, \ldots, \sigma_p^\dagger)$ with $\sigma_i^\dagger = \sigma_i^{-1}$ if $\sigma_i \neq 0$, $0^\dagger = 0$.

Proof. See [2] and [4].

LEMMA 2. *Let A be an m×n matrix.*

i) If rank $A < \min(m,n)$ then $X_k = 0$ is an unstable solution of (9).

ii) If rank $A < m$ then $T_k = 0$ is an unstable solution of (10).

Proof. It is sufficient to show that the linearized equations are unstable so the last terms of (9) and (10) are dropped

$$X_{k+1} = A_1 X_k + X_k A_2 \tag{17}$$

$$T_{k+1} = A_2 T_k + T_k A_2 \tag{18}$$

Let $A = U\Sigma V^T$, $X'_k = V^T X_k U$, $T'_k = U^T T_k U$.

Then from (11), (12), (16), (17), (18) one obtains

$$X'_{k+1} = D_1 X'_k + X'_k D_2 \tag{19}$$

$$T'_{k+1} = D_2 T'_k + T'_k D_2 \tag{20}$$

$$D_1 = I - \Sigma^{\dagger}\Sigma \tag{21}$$

$$D_2 = I - \Sigma\Sigma^{\dagger} \tag{22}$$

The diagonal matrices D_1 and D_2 have partly the same eigenvalues namely

$$\lambda_i = 1 - \sigma_i\sigma_i^{\dagger} = 1 - \sigma_i^{\dagger}\sigma_i = \begin{cases} 1 & \text{if} \quad \sigma_i = 0 \\ 0 & \text{if} \quad \sigma_i \neq 0 \end{cases}$$

According to the assumption i) there is at least one $\lambda_i = 1$, say $\lambda_r = 1$. The solution of (19) is

$$X'_k = \sum_{n=0}^{k} \begin{pmatrix} k \\ n \end{pmatrix} D_1^n \; X'_0 \; D_2^{k-n} \tag{23}$$

The element $(X'_k)_{rr}$ becomes

$$(X'_k)_{rr} = \sum_{n=0}^{k} \begin{pmatrix} k \\ n \end{pmatrix} (X'_0)_{rr} = 2^k (X'_0)_{rr}$$

and the unstability of (19) and consequently (9) is proven.

Noting that the assumption ii) implies that D_2 has an eigenvalue equal to 1, the proof of the second part of the lemma is quite analogous.

Q.E.D.

LEMMA 3.

i) *If all eigenvalues of X_0A have magnitudes less than one then $X_k = 0$ is an asymptotically stable solution of*

$$X_{k+1} = X_k A X_k \tag{24}$$

ii) *If all eigenvalues of T_0 have magnitudes less than one then $T_k = 0$ is an asymptotically stable solution of*

$$T_{k+1} = T_k^2 \tag{25}$$

Proof. Simple calculations (together with the assumptions) give

$$X_k = X_0(AX_0)^{2^k-1} = (X_0A)^{2^k-1}X_0 \to 0, \; k \to \infty$$

$$T_k = (T_0)^{2^k} \to 0, \; k \to \infty$$

Q.E.D.

LEMMA 4. *Let A be an m×n matrix. If all eigenvalues of X_0A have magnitudes less than one, and rank $A = n$ then $X_k = 0$ is a stable but not asymptotically stable solution of (9).*

Proof. The assumption implies $A_1 = 0$.

Introduce the sequences R_k and S_k by

$$R_k = X_k A_2 \tag{26}$$

$$S_k = X_k(I-A_2) \tag{27}$$

so $X_k = R_k + S_k$.

Equations (9), (12), (26) give

$$S_{k+1} = S_kAS_k \qquad (28)$$

with $S_0A \equiv X_0AA^{\dagger}A = X_0A$. Lemma 3 implies that the solution

$$S_k = (S_0A)^{2^k-1}S_0 \qquad (29)$$

is bounded and tends to zero as k → ∞. Equations (9), (12), (26), (27) give

$$R_{k+1} = (I+S_kA)R_k \qquad (30)$$

with the solution

$$R_k = (I+S_{k-1}A)(I+S_{k-2}A) \ldots (I+S_0A)R_0 \qquad (31)$$

Insert (29) and let $F = S_0A = X_0A$

$$R_k = (I+F^{2^{k-1}})(I+F^{2^{k-2}}) \ldots (I+F)R_0$$

or

$$R_k = (I-F)^{-1}(I-F^{2^k})R_0 \qquad (32)$$

The assumptions imply that the inverse exists. Thus the sequence R_k is bounded but does not tend to zero as k → ∞. The same is true about $X_k = R_k + S_k$.

Q.E.D.

6. MAIN RESULT.

Lemma 2 - 4 are now summarized in the following:

THEOREM. *Let A be an m×n matrix. Assume that (3) holds.*

i) *If rank A < min(m,n) then $X_k = 0$ is an unstable solution of (9).*

ii) *If rank A = min(m,n), $m \neq n$ then $X_k = 0$ is a stable, but not asymptotically stable, solution of (9).*

iii) *If rank A = m = n then $X_k = 0$ is an asymptotically stable solution of (9).*

iv) *If rank A < m then $T_k = 0$ is an unstable solution of (10).*

v) *If rank A = m then $T_k = 0$ is an asymptotically stable solution of (10).*

Proof. Cases i) and iv) follow immediately from Lemma 2.

Case ii) follows from Lemma 4 if all eigenvalues of X_0A have magnitudes less than one. The assumption rank A = $n(A^{\dagger}A = I)$ is easily transformed to the other possible case rank A = m by transposing (9) and (12). (13) gives

$$| \lambda_i(X_0A) | = | \lambda_i(A^{\dagger}A-\alpha A^TA) | = | \lambda_i(I-\alpha A^TA) | =$$

$$= | 1-\alpha\sigma_i^2 | < 1 \qquad (33)$$

according to (3). A *small* perturbation of (13) does not change this fact.

The cases iii) and v) follow by the same argument.

Q.E.D.

Remark. The conclusion of the theorem is that the algorithms (1), (2) cannot be used if rank A < min(m,n). If rank A = min(m,n) an ordinary QR algorithm may be used instead.

7. NUMERICAL EXAMPLES.

Some examples illustrating the properties of the algorithm are given below. Computations were carried out on a Univac 1108. The relative precision is about $1.5 \cdot 10^{-8}$. In accordance with [1] α was chosen as

$$\alpha = \frac{1}{\max\limits_{1\leq i\leq m} \sum\limits_{j=1}^{m} | b_{ij} |} , \quad B = AA^T \tag{34}$$

Observe that all figures have a logaritmic scale.

In the following the condition number of A denotes the quantity

$$\text{cond } A = || A || \cdot || A^{\dagger} || \tag{35}$$

with

$$|| A || = \sup_{|| x ||_2 = 1} || Ax ||_2 \tag{36}$$

From Lemma 1 one easily obtains

$$\text{cond } A = \frac{\max \sigma(A)}{\min\limits_{\sigma \neq 0} \sigma(A)} \tag{37}$$

Example 1

The matrix is

$$A(\varepsilon) = \begin{bmatrix} 0.5+\varepsilon & 0.5 & 0.5 \\ 0.5 & 0.5 & 0.5 \end{bmatrix}$$

with $\varepsilon = 10^{-3}$. In the beginning Y_k is very close to $[A(\varepsilon=0)]^{\dagger}$. In this example the convergence is slow. The condition number of A is $2 \cdot 10^3$.

Example 2

The matrix is

$$A = \begin{bmatrix} 1 & 2 & 3 & 6 & 0 \\ 2 & 0 & 1 & 3 & 1 \\ 1 & 3 & 2 & 6 & 2 \\ 0 & 1 & 1 & 2 & 0 \\ 1 & 5 & 0 & 6 & 6 \end{bmatrix}$$

Rank A = 3 and cond A ≈ 7. The numerical solution is unstable in accordance with the theorem. In the interval $10 < k < 25$ the matrices Y_k and Z_k do not change very much. In this interval they also differ very little from the true solution.

Example 3

The matrix is the inverse of the Hilbert matrix of order 6. The reason for choosing the inverse is given in [2]. The representation of the Hilbert matrix causes necessarily roundoff errors which will give a rather great deviation of the inverse.

The theorem implies stability in this case which seems to contradict the numerical result. The necessary relation of the singular values (33) hold $| 1-\alpha\sigma_i^2 | < 1$. Introduce

$$\varepsilon = \min_{\alpha} \max_{\sigma_i} 1 - | 1-\alpha\sigma_i^2 | \tag{38}$$

If A non-singular

$$\varepsilon = \frac{2}{1 + (\text{cond } A)^2} \tag{39}$$

For practical use it is reasonable to require $\varepsilon \geqslant \varepsilon_0$, the relative precision. Thus

$$\text{cond } A \leqslant \sqrt{\frac{2}{\varepsilon_0}} \tag{40}$$

is an upper bound of cond A.

For a non-singular matrix such an upper bound ought to be of the form

$$\text{cond } A \leqslant \frac{k}{\varepsilon_0} \tag{41}$$

for good methods.

In the actual example cond A $\approx 1 \cdot 10^7$ but $\sqrt{2/\varepsilon_0} \approx$ $\approx 1 \cdot 10^4$.

Example 4

The last column of the matrix of Example 3 was substituted by a linear combination of the first five columns.

In contrast to Example 2 there is no interval of k where Y_k and Z_k are almost constant and near the true solution.

8. SUPPLEMENTARY COMMENTS.

What information do the numerical examples yield relative to the questionsof Section 3?

The determination of rank A may fail in the unstable case. Only in special examples rank A < min(m,n) is consistent with stability. So it is in the example of [1].

$$a_{ij} = \frac{1}{10}\ ,\ i = j = 1, \ldots, 10$$

It can be shown that including the effect of roundoff errors $Y_k = AP_k$ with some matrices P_k, which implies $X_k' = 0$ in (19).

Example 1 shows that ill-conditioned problems require many iterations. The same is true if α is small.

For the matrix of Example 3 computing time and accuracy were compared with different algorithms. SVD and LSQ are based on the singular values decomposition. FORTRAN versions of the Algol programs given in [3] were used.

LSQR is an ordinary QR factorization of A, that is

$$A = Q\begin{bmatrix} R \\ \hdashline 0 \end{bmatrix}$$

Q orthogonal and R upper triangular [3]. Then

$$A^{\dagger} = [R^{-1} \,\vdots\, 0]Q^T$$

This method requires rank A = min(m,n).

Algorithm	Time in ms	Accuracy
SVD	44	$2\cdot10^{-2}$
LSQ	20	$2\cdot10^{-2}$
LSQR	7	$2\cdot10^{-2}$
Ben Israel & Cohen No. of iterations		
24	98	0.7
27	110	$1\cdot10^{-2}$
30	122	$2\cdot10^{-3}$
60	242	$4\cdot10^{-3}$

REFERENCES.

[1] A. Ben-Israel - D. Cohen: On Iterative Computation of Generalized Inverses and Associated Projections, SIAM J. Numer. Anal., 3 (1966), pp. 410-419.

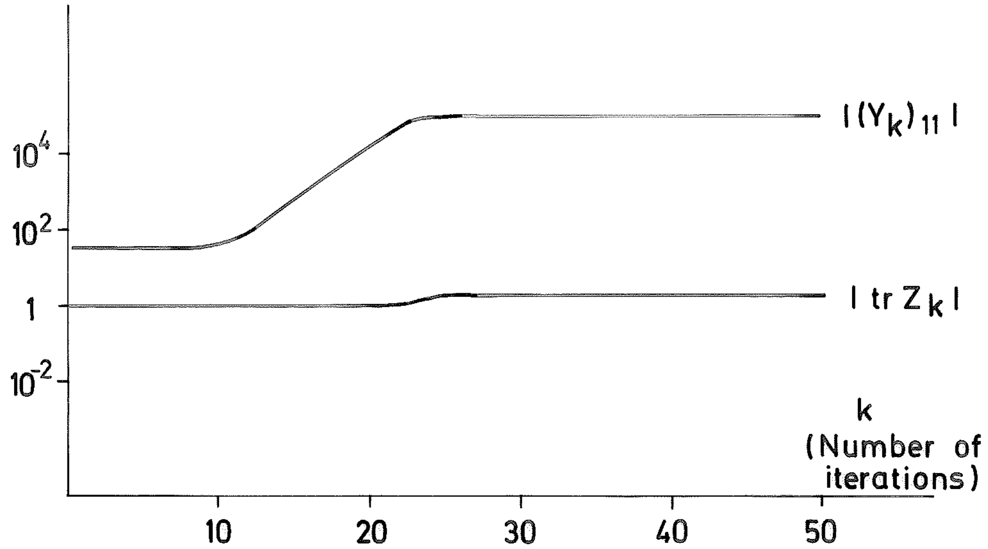
[2] G. Forsythe - C. Moler: Computer Solution of Linear Algebraic Systems, Prentice-Hall, Englewood Cliff, N.J. (1967).

[3] G.H. Golub: Numerical Methods for Solving Linear Least Squares Problems, Num. Math., 7 (1965), pp. 206-216.

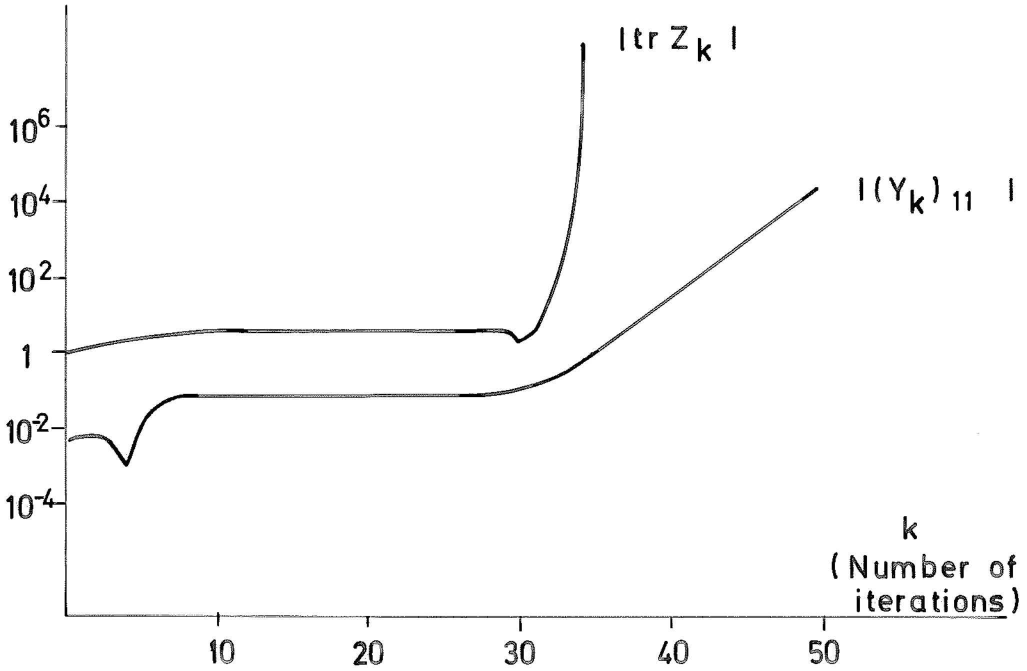
[4] G.H. Golub - C. Reinsch: Singular Value Decomposition and Least Squares Solution, Tech. Rep. No. CS 133, Comp. Sci. Dep., Stanford Univ. (1969).

[5] G. Peter - J.H. Wilkinsson: The Least Squares Problem and Pseudo-Inverses, The Comp. J., 13 (1970), pp. 309-316.

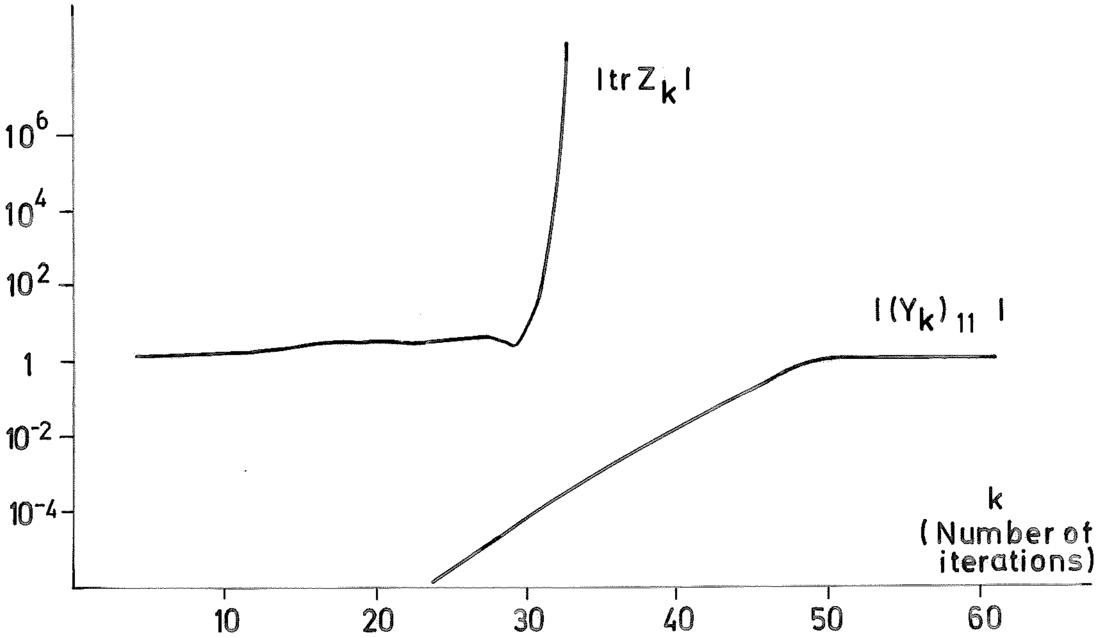
[6] T. Söderström: Notes on Pseudoinverses. Application to Identification, Report 7003, Div. of Aut. Control, Lund Inst. of Techn. (1970).



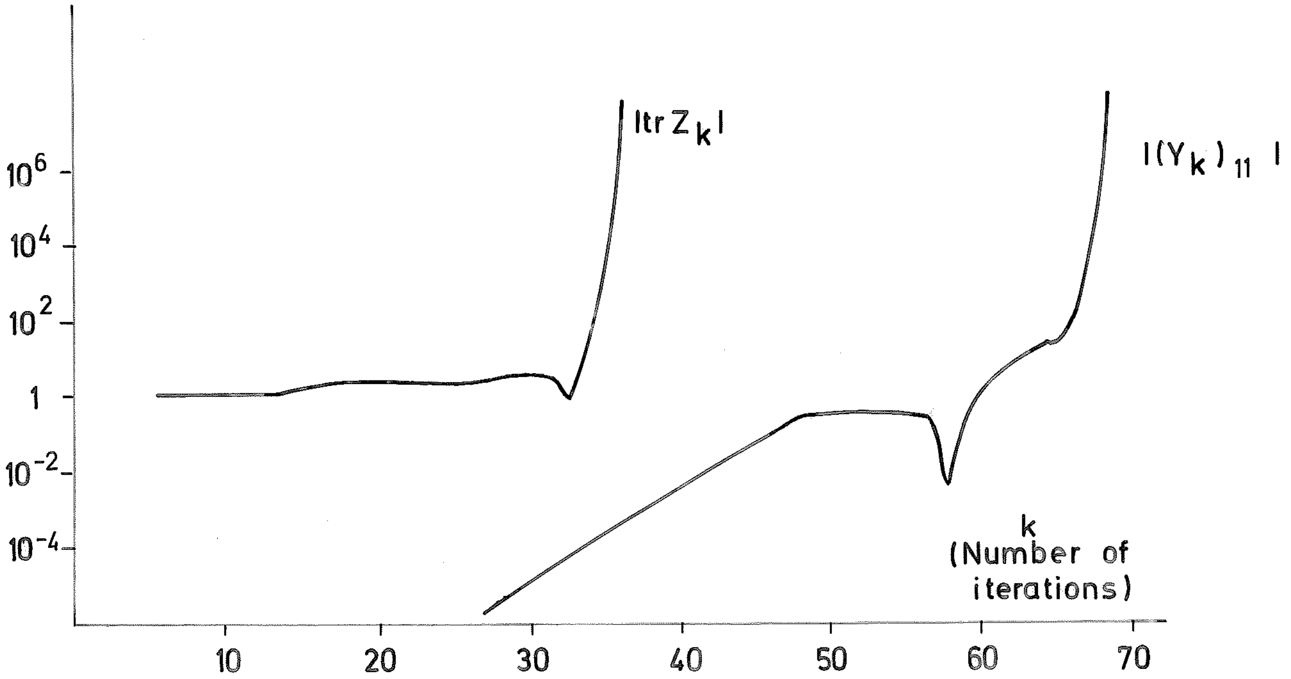
Example 1



Example 2



Example 3



Example 4