86-11C

PROCEEDINGS OF THE 25TH IEEE CONFERENCE ON DECISION AND CONTROL DECEMBER 10—12, 1986



Atheneum Intercontinental Hotel
Athens, Greece



VOLUME 3 OF 3

86CH2344-0

The 25th IEEE Conference on Decision and Control

Sponsoring Organization
IEEE Control Systems Society

Cooperating Organizations
Society for Industrial and Applied Mathematics (SIAM)
Operations Research Society of America (ORSA)

Supporting Organizations
The Greek Ministry of Culture
The National Technical University of Athens
Technical Chamber of Greece

and with the support and cooperation of
Institut National de Recherche Informatique et Automatique (INRIA)

Operating Committee

GENERAL CHAIRMAN
Professor Anthony Ephremides, *Electrical Engineering Dept., University of Maryland*

CO-CHAIRMAN
Professor Spyros Tzafestas, *National Technical University, Zografou, Athens*

PROGRAM CHAIRMAN
Professor H. Vincent Poor, *Coordinated Science Laboratory, University of Illinois*

FINANCE
Eyad Abed, *University of Maryland*
A. Kossidas, *SELETE, Marousi, Athens*

PUBLICATIONS
Sergio Verdu, *Princeton University*

PUBLICITY
Evaggelos Geraniotis, *University of Maryland*
A. Papakonstantinou, *National Technical University, Athens*

REGISTRATION
Andre Tits, *University of Maryland*
G. Stassinopoulos, *National Technical University, Athens*

EXHIBITS
Al Andry, *Teledyne Solid State, California*
E. Magirou (Greece), *Anotati Emporiki School of Athens*

LOCAL ARRANGEMENTS
P. Paraskevopoulos, *National Technical University, Athens*

A GEOMETRICAL APPROACH TO THE MAXIMAL CORANK PROBLEM IN THE ANALYSIS OF LINEAR RELATIONS.

Bart De Moor*;Joos Vandewalle.

* research sponsored by the IWONL.

ESAT-Katholieke Universiteit Leuven.
Kardinaal Mercierlaan 94 / 3030 Heverlee/ Belgium /tel 016 220931
telex 25941 elekul.

Abstract.

In this paper, a novel approach is provided to an important but unsolved mathematical problem that occurs in a wide variety of applications: Given a symmetric positive definite nxn matrix Σ, determine all diagonal nonnegative matrices $\tilde{\Sigma}$ so that the difference matrix $\hat{\Sigma} = \Sigma - \tilde{\Sigma}$ is nonnegative definite and its rank is minimal.
In this paper, we explore the geometrical properties of the solution vectors x satisfying $\hat{\Sigma}.x=0$. New concepts such as orthant and null invariance are introduced. The results in this paper are of key importance in the analysis of noisy linear equations (factor analysis. They will lead to a complete geometrical characterization of the solution set, which will be described in a forthcoming paper.

Keywords: Sets of linear equations, maximal corank, least squares, non-uniqueness, factor analysis.

1. Introduction.

The central idea of this work is the 'non-uniqueness principle' (KAL), which states that the solution of a noisy identification problem is intrinsically non-unique: Any uncertainty in the data implies that no particular model can be identified with certainty. Although this point of view seems quite natural, it contradicts what could be called the 'classical' approach to the identification problem, where in a lot of cases mathematical artefacts are introduced in order to ensure the uniqueness of the model (e.g. minimization of norms in least squares and orthogonal regression approaches, maximum likelihood,...).
The main goal of this paper is to introduce the mathematical tools to analyse the uncertainty principle for the basic modelling problem of the identification of static linear ..ations from noisy data. This problem is not as restricted as .. may appear at first sight. A lot of identification and dynamic modelling problems are formulated in terms of static linear equations.
A nice feature of the geometrical approach here presented is that only elementary linear algebra is needed. The complete geometrical characterization of the global solution set in terms of the geometrical properties derived in this paper, will be provided in a forthcoming paper.
This paper is organized as follows: In section 2 some fundamental assumptions are formulated together with the problem statement. In section 3 we proof an important lemma about the classical linear least squares solutions. Important original theorems, describing geometrical properties of the solution vectors are given in section 4. In section 5, an important noise space theorem is derived while in section 6, the properties of some special allowed orthant null invariant vectors are studied.
The following notations and abbreviations are used throughout:

A	mxn matrix with m>n	PD	Positive Definite
A^t	transpose	NND	Nonnegative definite
r(A)	rank of the matrix A	E(.)	Exspectation operator
cor(A)	corank of A = n-r(A)	ker(A)	kernel of matrix A
e_i	$(0...0\ 1\ 0...0)^t$		$=(x\in\mathbb{R}^n \mid A.x=0)$
I_n	square identity matrix	diag(±1)	sign matrix with ±1
$\hat{A}$	exact matrix		as diagonal elements
$\tilde{A}$	pure noise matrix	diag(σ_i)	diagonal matrix with
A>B	$a_{ij}>b_{ij}\ \forall\ i,j$		diagonal elements σ_i
$\mathbb{R},\mathbb{Z},\mathbb{N}$	real, integer, natural numbers	LLS	linear least squares
OI	Orthant Invariant		
ONI	Orthant Null Invariant		
AONI	Allowed Orthant Null Invariant		

2. Problem statement.

Let us consider a vector sequence $(a_i, i=1,m)$ of data vectors a_i in $\mathbb{R}^n$, stored as rows in a mxn matrix A. This matrix can be considered as m subsequent observations (measurements) on a process with n observed variables. It is assumed without loss of generality, that all variables have zero mean :

$$\sum_{i=1}^{m} a_{ij} = 0 \quad \text{for } j=1,..,n$$

The problem is to find all possible n-vectors x that describe a linear relation between the columns of A:

$$\underset{mxn}{A} \cdot \underset{nx1}{x} = \underset{mx1}{0} \qquad (2.1)$$

This is equivalent to finding all possible vectors x that satisfy:

$$\underset{nxn}{\Sigma} \cdot \underset{nx1}{x} = \underset{nx1}{0} \qquad (2.2)$$

where $\Sigma=E(A^t.A)$ is the covariance matrix of the vector process a_i. The problem of statistical sampling errors and the robustness of the solution scheme against it, will not be considered here. For the moment, this would be of no relevance to the purpose of this paper, which is, to clarify, by a geometrical approach, the link between the uncertainty in the data and the non-uniqueness of the solution.
Clearly, at least one solution to problems (2.1) and (2.2) exists if the linear algebraic rank condition is satisfied:

$$r(A) = r(\Sigma) < n \text{ or equivalently } cor(A) = cor(\Sigma) \geq 1 \qquad (2.3)$$

The number of linear independent solutions equals cor(Σ), the dimension of the kernel of Σ, which is the solution space.

However, in general, due to the noise, the measurement covariance matrix will satisfy cor(Σ)=0 and hence no trivial solution could be obtained directly from the data, which must first be modified to obtain a solution. But this requires some fundamental assumptions.

Assumption 1 : cor(Σ) = 0.

It states that, due to the noise, there are no trivial linear relations. It follows directly that Σ^{-1} exists. This inverse covariance will appear to be the key tool in the forthcoming analysis (see also [DEM2] for more results on Σ^{-1}).

Assumption 2 : The uncertainties are additive.

Each column of the data matrix A of (2.1) is the sum of an exact and a pure noise column:

$$A = \hat{A} + \tilde{A} \qquad (2.5)$$

In this paper, it is assumed that all data are noisy.

Assumption 3 : Each noise column of the pure noise matrix $\tilde{A}$ is independent of (uncorrelated with) each exact data column in the exact data matrix $\hat{A}$.

All questions of temporal correlations are disregarded. Although this may look a limitation of the scope of the paper, almost all 'classical' noise schemes implicitely contain a similar assumption. Moreover, as will be demonstrated, even under this restrictive assumption, the general solution is complicated and essentially non-unique.
Assumptions 2 and 3 lead to the following well known covariance formulation:

$$\Sigma = \hat{\Sigma} + \tilde{\Sigma} \qquad (2.6)$$

stating that the measurement covariance matrix Σ is the sum of an exact covariance matrix $\hat{\Sigma}$ and a noise covariance matrix $\tilde{\Sigma}$.

Assumption 4: The noise covariance matrix $\tilde{\Sigma}$ is NND diagonal.

This axiom is inevitable if the problem of identifying linear relations is to be well defined. Because if $\tilde{\Sigma}$ were not diagonal, there would exist residual linear relations between the uncertainties in addition to the linear relations already found in Σ-$\tilde{\Sigma}$. The noise covariance matrix $\tilde{\Sigma}$ could then be analyzed further, according to the solution scheme presented furtheron, leading to other linear relations.

The set of 4 fundamental assumptions is known as the 'FRISCH scheme' [KAL]. The problem statement now becomes:

Noise space problem formulation.

Given a nxn symmetric, PD matrix Σ with cor(Σ)=0.
Find all diagonal, NND matrices $\tilde{\Sigma}$ such that

1. $\hat{\Sigma} = \Sigma - \tilde{\Sigma}$
2. $\hat{\Sigma}$ is NND
3. cor($\hat{\Sigma}$) is maximal (2.7)

A vector x is called a solution vector if it satisfies:

$$\hat{\Sigma} \,.\, x = (\Sigma - \tilde{\Sigma}) \,.\, x = 0 \qquad (2.8)$$

An alternative problem formulation is:

Solution space problem formulation.

Given a nxn symmetric, PD matrix Σ with cor(Σ)=0.
Find all vectors x such that:

1. There exists a NND diagonal $\tilde{\Sigma}$ for which
   $\Sigma \,.\, x = \tilde{\Sigma} \,.\, x$
2. $\hat{\Sigma} = \Sigma - \tilde{\Sigma}$ is NND
3. cor($\hat{\Sigma}$) is maximal (2.9)

Obviously, the NND constraint for the difference matrix $\hat{\Sigma}$ is a necessary condition for a covariance matrix. But also the third condition of maximal corank is crucial. The problem of identifying linear relations is not well defined unless the corank is required to be maximal. Because, if the data Σ support c (=max cor(Σ)) different linear independent linear relations, then it will support infinitely many linear combinations of only c'<c of the c linear relations. Hence only the maximal corank is a identifiable invariant of the data.
Although the noise space formulation (2.7) does not differ much from the solution space problem formulation (2.9), the solution techniques are different. The main existing results for the noise space problem formulation were clarified by Kalman [KAL].
The results in this paper can be considered as the basic tools for the solution of the solution space problem.

3. Two classical solutions.

If the requirement of maximal corank is dropped from the problem formulation, some classical solutions with diagonal NND noise covariance matrix are easily derived.

3.1 The linear least squares solutions.

The following theorem is a remarkable but little known result.

Lemma 1 :
Given a nxn symmetric, PD matrix Σ with cor(Σ)=0. The columns of Σ^{-1} are the n possible least squares solutions of the set of linear equations with data matrix Σ, with for the i-th LLS solution s_i (i-th column of Σ^{-1}) the noise covariance matrix:

$$\tilde{\Sigma}_i = \mathrm{diag}(0,\ldots,0,1/s_{ii},0,\ldots,0) \qquad (\uparrow\ i)$$

where s_{ii} is the i-th diagonal element of $S=\Sigma^{-1}$ and

$$(\Sigma - \tilde{\Sigma}_i).s_i = 0 \text{ with } \Sigma - \tilde{\Sigma}_i \text{ NND.}$$

Proof: a/ s_i satisfies $(\Sigma - \tilde{\Sigma}_i).s_i=0$.
Because r(Σ)=n, Σ^{-1} exists. Let $s_i=(s_{1i}\ s_{2i}\ldots s_{ni})^t$ be the i-th column of Σ^{-1}. Then :

$$\Sigma.s_i=e_i=(0..0\ 1\ 0..0)^t$$

This can be written as:

$$(\ \Sigma - \mathrm{diag}(0..0\ 1/s_{ii}\ 0..0)\ ).s_i = 0$$

or

$$(\ \Sigma - \tilde{\Sigma}_i)\ .s_i = 0$$

Q.E.D.

b/ $\Sigma - \tilde{\Sigma}_i$ is NND.
Because Σ is PD, also Σ^{-1} is PD. Moreover, it is easily seen that:

$$(\Sigma - \tilde{\Sigma}_i).\Sigma^{-1}.(\Sigma - \tilde{\Sigma}_i) \text{ is NND.}$$

and it is easy to verify that $\tilde{\Sigma}_i.\Sigma^{-1}.\tilde{\Sigma}_i=\tilde{\Sigma}_i$ so that

$$(\Sigma - \tilde{\Sigma}_i).\Sigma^{-1}.(\Sigma - \tilde{\Sigma}_i) = \Sigma - \tilde{\Sigma}_i \text{ is NND.}$$

Because $r(\tilde{\Sigma}_i)=1$, and due to the diagonal structure of $\tilde{\Sigma}_i$, it follows that only measurements on the i-th variable are considered as noisy and only those are modified in order to obtain a solution. But this corresponds to the classical geometrical ('projection') interpretation of the mechanism of LLS (see also [DEM2]).

Q.E.D.

The inverse covariance matrix Σ^{-1} contains the n possible least squares solutions and also the least squares noise information $\sigma_i=1/s_{ii}$. From the preceeding proof it is obvious that the i-th LLS solution s_i can only be justified if only measurements on the i-th variable are noisy, all others being noiseless. In that case, the solution is unique. However, if all data are noisy, the LLS approach does not provide a satisfying answer to the problem. Clasically, this is then solved by modelling the noise covariance matrix as a multiple of the identity matrix.

3.2 The noise covariance matrix as a multiple of the identity matrix.

Lemma 2 :
Given a nxn symmetric, PD matrix Σ with $r(\Sigma)=n$ and with eigenvalue decomposition:

$$\Sigma = \sum_{i=1}^{n} d_i.v_i.v_i^t$$

where $d_i \geq d_{i+1}$ are the eigenvalues and v_i the eigenvectors.
Let $\tilde{\Sigma} = d_n.I_n$. Then :

1. $(\Sigma - \tilde{\Sigma}).v_n = 0$
2. $\Sigma - \tilde{\Sigma}$ is NND

Proof: Follows immediately from the properties of the eigenvalue decomposition of a PD matrix.

Q.E.D.

The implicit assumption here is that all variables are perturbed by the same amount of noise and that the variance of this noise equals the smallest eigenvalue. Again ,the solution v_n is unique if d_n is simple. Of course, this technique can provide satisfactory answers, if it is possible to scale the data columns of the matrix A so that they all have an equal noise variance σ. This assumes a perfect a priori knowledge of the noise characteristics which, in some cases, is readily available. Nevertheless, when such information is missing, the restriction of $\tilde{\Sigma}$ to a multiple of I_n is not justifiable. But obviously, this specific solution will belong to the general solution set of problem (2.9).

4. Geometrical properties of the solution vectors x .

In this section some important geometrical properties of the solution vectors are studied. Novel geometrical concepts such as orthant and null invariance are introduced. Those novel characterizations allow to formulate the main theorems that describe geometrically the solution set.

4.1. Corank n-r solutions.

The following theorem, which is one of classical linear algebra, suggests how solution vectors with zero components will play a crucial role in the determination of the maximal corank.

Theorem 1:
Given the exact NND matrix $\hat{\Sigma}$ with $r(\hat{\Sigma})=r < n$. The problem of finding the vectors x satisfying $\hat{\Sigma}.x=0$, always has $n-r=cor(\hat{\Sigma})$ linear independent solutions, each of which has at least n-r-1 zero components.

Proof: Because $r(\hat{\Sigma})=r < n$, there exists a nxn column-permutation matrix P so that $\hat{\Sigma}.P = (\underset{nxr}{\hat{\Sigma}_1} \quad \underset{nx(n-r)}{\hat{\Sigma}_2})$ with $r(\hat{\Sigma}_1)=r$.
If now x satisfies $\hat{\Sigma}.x=0$, and x is partitioned suitably, then $\hat{\Sigma}_1.x_1=-\hat{\Sigma}_2.x_2$ and using the pseudo-inverse $\hat{\Sigma}_1^+$ of $\hat{\Sigma}_1$, we have $x_1=-\hat{\Sigma}_1^+.\hat{\Sigma}_2.x_2$. Now take $x_2=e_i^t = (0...0\ 1\ 0...0)^t$ for i=1,...,n-r.

then the n-r columnvectors of $\begin{bmatrix} \\ -\Sigma_1^+.\Sigma_2 \\ \\ \end{bmatrix}$ $\begin{matrix} \\ \\ I_{n-r} \end{matrix}$

are n-r linear independent solutions of $\Sigma.x=0$ each with at least n-r-1 zero components.

Q.E.D.

The geometrical interpretation is that, if the corank is n-r, there are n-r linear independent solutions that lie in n-r different coordinate planes. This observation is important for the determination of the maximal corank, as will be demonstrated furtheron.

4.2. Orthant invariance.

The property of orthant invariance is inspired by the observation that, if x satisfies $(\Sigma-\tilde{\Sigma}).x=0$, then also $\Sigma.x=\tilde{\Sigma}.x$. Since $\tilde{\Sigma}$ is NND diagonal, the vector $\Sigma.x$ will have the same sign pattern as the vector x.

Definition 1: Given a nxn diagonal sign matrix E=diag(±1). A n-vector x with components x_i is said to belong to orthant E: $x \in E$ if $E.x_i \geq 0$, $\forall$ i=1,..,n.

Definition 2: Orthant Invariance.
Given a square nxn matrix H, a diagonal sign matrix E=diag(±1) and a vector $x \in E$. Then x is orthant invariant (OI) for the matrix H if from $x \in E$ it follows that $H.x \in E$.

We are now ready to formulate a first geometrical characterization of all solution vectors of problem (2.9):

Lemma 3:
Given a nxn symmetrix PD matrix Σ. All vectors x satisfying $(\Sigma - \tilde{\Sigma}).x=0$ or equivalently $\Sigma.x=\tilde{\Sigma}.x$ where $\tilde{\Sigma}$ is NND diagonal, are orthant invariant (OI) for Σ.

Proof: Follows immediately from definition 1 and 2.

Q.E.D.

So all solution vectors of problem (2.9) must be OI.
A second important property is what we will call ' the convexity of the orthant invariance property in an orthant ',

Lemma 4: Convexity of orthant invariance.
Let $x_1 \in E$ and $x_2 \in E$ be two OI vectors for a square nxn matrix H. Then any convex combination x of x_1 and x_2 will be OI for H.
So if $x = \alpha.x_1+(1-\alpha).x_2$ with $0 \leq \alpha \leq 1$ then $x \in E$ and it follows that $H.x \in E$.

Proof: Trivial.

This property will appear to be extremely important for the geometrical description of the global solution set in one orthant.
We now define another important concept.

4.3.Null invariance.

Definition 3: Null invariance componentwise.
Given a square nxn matrix H and a n-vector x with $x_i=0$ for some $i \in \{1,..,n\}$. Let y=H.x. Then x is null invariant (NI) for component i : if from $x_i=0$,it follows that $y_i=(H.x)_i=0$.

The geometrical interpretation is that the vector $\Sigma.x$ of a null invariant x with a zero component has at least the same pattern of zeros as the vector x and hence lies in the same coordinate plane as x.
By convention, if the vector x has no zero components, it will always be considered as null invariant.
From this definition and convention, the second geometrical characterization of all solution vectors of problem (2.9) is easily derived:

Lemma 5:
Given a nxn symmetric PD matrix Σ. All vectors x satisfying $(\Sigma - \tilde{\Sigma}).x=0$ for some NND diagonal matrix $\tilde{\Sigma}$ are null invariant (NI).

Proof: If $x_i=0$, obviously $(\tilde{\Sigma}.x)_i=0$ hence $(\Sigma.x)_i=0$. Q.E.D.

The following lemma clarifies the crucial importance of the previously introduced concepts of orthant and null invariance.

Lemma 6: Existence of a diagonal matrix.
Given a nxn symmetric PD matrix Σ. For every ONI vector x of Σ there exists a NND diagonal matrix $\tilde{\Sigma}$ such that $\Sigma.x=\tilde{\Sigma}.x$.

Proof: Let $y=\Sigma.x$. Define $\tilde{\sigma}_i$ via $y_i=\tilde{\sigma}_i.x_i$.
If $x_i\neq 0$ then $\tilde{\sigma}_i=y_i/x_i\geq 0$ because x is OI.
If $x_i=0$ then $\tilde{\sigma}_i$ can be any nonnegative real number because x is NI, hence $y_i=0$.
For a ONI vector x, the corresponding NND diagonal $\tilde{\Sigma}$ then becomes $\Sigma=\mathrm{diag}(\tilde{\sigma}_i)$.

Q.E.D.

Remark that if some components of the ONI vector x are zero, the diagonal matrix $\tilde{\Sigma}=\mathrm{diag}(\tilde{\sigma}_i)$ is NOT unique.

4.4. Allowed Orthant Null Invariant vectors.

In lemma 6, the existence of a diagonal matrix associated to a ONI vector x was proved. However, nothing was said about the nonnegativedefiniteness of the difference $\Sigma - \tilde{\Sigma}$.

Definition 4: Given a nxn PD symmetric matrix Σ, a ONI vector x and a corresponding NND diagonal matrix $\tilde{\Sigma}$ (lemma 6). The vector x is allowed if $\Sigma - \tilde{\Sigma}$ is NND.

The complete characterization of the global solution set of the solution space problem, is provided in the following theorems:

Theorem 2 : Geometrical description of the solution set.

Given a nxn symmetric PD matrix Σ with cor(Σ)=0. The vectors x, satisfying $(\Sigma - \tilde{\Sigma}).x=0$, where $\tilde{\Sigma}$ is NND diagonal and $(\Sigma - \tilde{\Sigma})$ is NND are allowed, orthant, null invariant (AONI) vectors for Σ.

Corollary: The least squares solutions, described in lemma 1, are AONI vectors.

Theorem 3 : Maximal corank.

The maximal corank (defined in problem (2.9)) is n-r if there do not exist more than n-r linear independent AONI-vectors, each with n-r-1 zero components.

5. A property of the diagonal noise covariance matrix.

In lemma 6 we showed how to find for a ONI vector x a NND diagonal matrix $\tilde{\Sigma}$ such that $\Sigma.x=\tilde{\Sigma}.x$.
The following theorem determines the maximal noise values of the elements of the noise covariance $\tilde{\Sigma}$. This theorem will be important for the proofs of the results in section 6.

Theorem 4: Maximal noise theorem.
Given a nxn symmetrix PD matrix Σ. Let $S=\Sigma^{-1}$ with elements s_{ij}. Define $\tilde{\Sigma} = \mathrm{diag}(\tilde{\sigma}_i)$ .. If any of the elements $\tilde{\sigma}_i > 1/s_{ii}$, then $\Sigma - \tilde{\Sigma}$ has at least one negative eigenvalue and is NOT NND.

Proof: Let $\Sigma = \sum_{i=1}^{n} p_i.\Sigma_i$ where $\tilde{\Sigma}_i=\mathrm{diag}(0..1/s_{ii}..0)$ are the LLS noise covariances (lemma 1) and let $p_i\geq 0$, $\forall$ i. Let

$s_k=(s_{1k}...s_{nk})^t$ be the k-th column of $S=\Sigma^{-1}$.

Consider the quadratic form :

$$s_k^t.(\Sigma - \tilde{\Sigma}).s_k = s_k^t.(\Sigma - \sum_{i=1}^{n} p_i.\tilde{\Sigma}_i).s_k$$
$$= s_{kk} - \sum_{i=1}^{n} p_i^2.s_{ik}^2/s_{ii}$$
$$= (1 - p_k).s_{kk} - \sum_{\substack{i=1 \\ i\neq k}}^{n} p_i.s_{ik}^2/s_{ii}.$$

Since $p_i^2\geq 0$,$s_{ik}^2\geq 0$ and $s_{ii}>0$, this quadratic form is certainly negative if $p_k>1$. In that case $\tilde{\sigma}_k=p_k/s_{kk}>1/s_{kk}$ and $\Sigma-\tilde{\Sigma}$ is NOT NND.

Q.E.D.

Noise covariance matrices $\tilde{\Sigma}$ with at least one diagonal element larger than the corresponding least squares noise (lemma 1), can never be a solution to problem (2.7). However,this condition of theorem 4 is necessary but not sufficient : There exist matrices $\tilde{\Sigma}=\mathrm{diag}(\tilde{\sigma}_i)$ where all $\tilde{\sigma}_i<1/s_{ii}$ but where $\Sigma-\tilde{\Sigma}$ is not NND.

6. Properties of AONI-vectors.

In this section, we will explore the properties of some special AONI vectors.

The technique to find all OI vectors for a given nxn matrix Σ of full rank n is combinatorial in the sense that for each orthant $E=\mathrm{diag}(\pm 1)$ the set of OI vectors will be determined. Only half of the orthants has to be considered, since the other half is formed by the opposite orthants. Hence $2^n/2=2^{n-1}$ orthants are to be checked. The general procedure to compute all OI vectors in an orthant will not be described here. This will be the subject of a forthcoming paper. The central idea of the general procedure arises from the following observations:

Lemma 7:
The set of vectors x such that $\Sigma.x \in E$ is described by $x = \Sigma^{-1}.u$ where $u \in E$.

Proof: Trivial.

Now, if $x = \Sigma^{-1}.u$ with $u \in E$, then $\Sigma.x \in E$ and a condition for x to be OI is clearly $x \in E$. This can also be reformulated towards the nonnegative orthant:
If $u \in E$, then $E.u \in I_n$ and x is OI if $E.x = (E.\Sigma^{-1}.E).(E.u) \in I_n$ where now E.u is a nonnegative vector which makes nonnegative linear combinations of the columns of $(E.\Sigma^{-1}.E)$. When not all the columns of $(E.\Sigma^{-1}.E)$ are in the nonnegative orthant, there exist nonnegative linear combinations E.u so that $E.x=(E.\Sigma^{-1}.E).(E.u)$ s negative components. These vectors x are obviously NOT OI.
This means that the convex simplicial set, generated by the columns of $(E.\Sigma^{-1}.E)$ must be restricted to the nonnegative orthant. However the procedure for this restriction will not be described here but will be the subject of a forthcoming paper. Instead, we will concentrate on the first step of this general procedure and explore the properties of the vectors of the first step.
The starting point is the computation for each sign matrix E, of all convex combinations x of pairs of columns of $E.\Sigma^{-1}.E$ such that x has at least one zero component. For the remaining part of this paper, we will concentrate upon the properties of vectors that are convex combinations of columns of $E.\Sigma^{-1}.E$ and have at least one zero component.

Let s_i and s_j be two columns of the matrix $E.\Sigma^{-1}.E$.
Let x be a convex combination of s_i and s_j :
$$x = \alpha.s_i+(1-\alpha).s_j = \alpha.(E.\Sigma^{-1}.E).e_i+(1-\alpha).(E.\Sigma^{-1}.E).e_j$$
with $0 < \alpha < 1$ such that x_k, the k-th component of x is zero:
(For $\alpha=0$ and $\alpha=1$, we simply obtain the least squares vectors of lemma 1.)

ence : $x_k = \alpha.s_{ki}+(1-\alpha).s_{kj}=0$ so that $\alpha = \alpha_k = s_{kj}/(s_{kj}-s_{ki})$

For vectors x of this form it is not difficult to check the following property:

Lemma 8:
$E.\Sigma.E.\ x = E.\Sigma.E.(\alpha.s_i+(1-\alpha).s_j) = \alpha.e_i+(1-\alpha).e_j$

Proof: Trivial.

We will use this property to proof the following theorems:

Theorem 5: Null invariance.
Let x be a convex combination of columns s_i and s_j of $E.\Sigma^{-1}.E$ so that $x_k=0$. Then
x is NI for component k if $k \neq i$ and $k \neq j$.

Proof: Follows immediately from the definition of NI and lemma 8.

Theorem 6: Orthant Invariance.
Let x be a convex combination of columns s_i and s_j of $E.\Sigma^{-1}.E$ so that $x_k=0$. Then
E.x is OI for Σ if $(s_{kj}/s_{ki}) \leq (s_{ij}/s_{ii})$
and
$(s_{ki}/s_{kj}) \leq (s_{ii}/s_{ij})$

Proof: From lemma 8 and the definition of OI it follows that x is OI for $E.\Sigma^{-1}.E$ if $x_i \geq 0$ and $x_j \geq 0$.
$x_i = \alpha.s_{ii}+(1-\alpha).s_{ij} \geq 0$ if $\alpha/(\alpha-1) \leq s_{ij}/s_{ii}$
and
$x_j = \alpha.s_{ji}+(1-\alpha).s_{jj} \geq 0$ if $(\alpha-1)/\alpha \leq s_{ji}/s_{jj}$

Because $x_k=\alpha.s_{ki}+(1-\alpha).s_{kj}=0$, $\alpha/(\alpha-1)=s_{kj}/s_{ki}$.
Combination of these three inequalities, leads to the result stated. Moreover, if x is OI for $E.\Sigma.E$, then E.x is OI for Σ.

Q.E.D.

Theorem 7: ONI vectors that are NOT allowed.
Let x be a convex combination of columns s_i and s_j of $E.\Sigma^{-1}.E$ so that $x_k=0$. Let x be an ONI vector for $E.\Sigma^{-1}.E$.
Then E.x is NOT allowed if $s_{ij} < 0$.

Proof: Since x is ONI for $E.\Sigma^{-1}.E$, it follows from lemma 6 that there exists a NND diagonal matrix $\tilde{\Sigma}$ such that $\Sigma.(E.x)=\tilde{\Sigma}.(E.x)$.
Now define $y=\Sigma.(E.x)=\alpha.(E.e_i)+(1-\alpha).(E.e_j)$ (lemma 8) .
Define $\tilde{\Sigma} = \mathrm{diag}(\tilde{\sigma}_k)$ as:
$\tilde{\sigma}_k = 0$ for $k \neq i$ and $k \neq j$.
For k=i and k=j we have the following results:

$\tilde{\sigma}_i$	$\tilde{\sigma}_j$
$x_i=\alpha.s_{ii}+(1-\alpha).s_{ij}>0$	$x_j=\alpha.s_{ji}+(1-\alpha).s_{jj}>0$
$\tilde{\sigma}_i=y_i/x_i$	$\tilde{\sigma}_j=y_j/x_j$
$= \alpha/(\alpha.s_{ii}+(1-\alpha).s_{ij})$	$=(1-\alpha)/(\alpha.s_{ji}+(1-\alpha).s_{jj})$
But $s_{ij} < 0$ so that	But $s_{ij} < 0$ so that
$\alpha.s_{ii}+(1-\alpha).s_{ij}<\alpha.s_{ii}$	$\alpha.s_{ji}+(1-\alpha).s_{jj}<(1-\alpha).s_{jj}$
and hence	and hence
$\tilde{\sigma}_i=\alpha/(\alpha.s_{ii}+(1-\alpha).s_{ij})>1/s_{ii}$	$\tilde{\sigma}_j=(1-\alpha)/(\alpha.s_{ji}+(1-\alpha).s_{jj})>1/s_{jj}$

Hence both $\tilde{\sigma}_i$ and $\tilde{\sigma}_j$ violate the maximal noise theorem and x is NOT allowed.

Q.E.D.

Theorem 8: ONI vectors that are allowed.
Let x be a convex combination of the columns s_i and s_j of $E.\Sigma^{-1}.E$ so that $x_k=0$ and $\alpha=s_{kj}/(s_{kj}-s_{ki})$. Let x be ONI for $E.\Sigma^{-1}.E$. Then x is allowed if $s_{ij} > 0$.

In order to proof theorem 8 we need some classical lemma's, the proof of which will be omitted for brevity.

Lemma 9:
Given a nxn symmetric PD matrix Σ partitioned as
$$\Sigma = \begin{bmatrix} \Sigma_{11} & \Sigma_{12} \\ t & \\ \Sigma_{12} & \Sigma_{22} \end{bmatrix}$$ where Σ_{ij} is pxp (p<n)
then Σ_{11} and Σ_{22} are also PD.

Proof: Follows immediately from the interlacing property of the eigenvalues of submatrices of symmetric matrices [GAN].

Q.E.D.

Lemma 10: Block matrix inversion lemma.
If A^{-1} and B^{-1} exist, then

$$\begin{bmatrix} A & D \\ C & B \end{bmatrix}^{-1} = \begin{bmatrix} (A-D.B^{-1}.C)^{-1} & -A^{-1}.D.(B-C.A^{-1}.D)^{-1} \\ -(B-C.A^{-1}.D)^{-1}.C.A^{-1} & (B-C.A^{-1}.D)^{-1} \end{bmatrix}$$

Proof: see [GAN].

Lemma 11:
Define $\tilde{\Sigma} = \begin{bmatrix} \tilde{\Sigma}_1 & 0 \\ 0 & 0 \end{bmatrix}$ where $\tilde{\Sigma}_1$ is NND diagonal of dimension n-r+1.

Let $cor(\Sigma - \tilde{\Sigma}) = n-r$ and partition $S = \Sigma^{-1} = \begin{bmatrix} S_{11} & S_{12} \\ t & \\ S_{12} & S_{22} \end{bmatrix}$

and $\Sigma = \begin{bmatrix} \Sigma_{11} & \Sigma_{12} \\ t & \\ \Sigma_{12} & \Sigma_{22} \end{bmatrix}$

where S_{11} and Σ_{11} are square (n-r+1)x(n-r+1)
Then $cor(\Sigma - \tilde{\Sigma}) = cor(S_{11}^{-1} - \tilde{\Sigma}_1) = n-r$.

Proof : Omitted for brevity.

Lemma 12 :

If $P = \begin{bmatrix} P_{11} & P_{12} \\ t & \\ P_{12} & P_{22} \end{bmatrix}$ is a symmetric 2x2 block matrix with P_{22} PD.

then

$$P \text{ is PD (NND)} \iff (P_{11} - P_{12}.P_{22}^{-1}.P_{12}^{t}) \text{ is PD (NND)}$$

Proof : See [KAL, p.154].

Lemma 13 :

Given the nxn PD matrix Σ partitioned as $\Sigma = \begin{bmatrix} \Sigma_{11} & \Sigma_{12} \\ t & \\ \Sigma_{12} & \Sigma_{22} \end{bmatrix}$

and $\tilde{\Sigma} = \begin{bmatrix} \tilde{\Sigma}_1 & 0 \\ 0 & 0 \end{bmatrix}$ where $\tilde{\Sigma}_1$ is NND diagonal of dimension n-r+1 so that $cor(\Sigma - \tilde{\Sigma}_1)=n-r$. Let $S = \Sigma^{-1} = \begin{bmatrix} S_{11} & S_{12} \\ t & \\ S_{12} & S_{22} \end{bmatrix}$

and both Σ_{11} and S_{11} are square (n-r+1)x(n-r+1).
Then :

$$[S_{11}^{-1} - \tilde{\Sigma}_1] \text{ is NND} \iff \Sigma - \tilde{\Sigma} \text{ is NND} .$$

Proof : Follows from lemma 9,10,11,12.

ith lemma 9 to 13 we are now ready to prove theorem 8.

Proof of theorem 8: Because x is ONI for $E.\Sigma^{-1}.E$ there exists a diagonal matrix $\tilde{\Sigma}$ such that $\Sigma.(E.x)=\tilde{\Sigma}.(E.x)$ (lemma 6).
Define $\tilde{\Sigma} = diag(\tilde{\sigma}_k)$ as:
$\tilde{\sigma}_k=0$ if $k\neq i$ and $k\neq j$. For $k=i$, $\tilde{\sigma}_i=\alpha/(\alpha.s_{ii}+(1-\alpha).s_{ij})$.
For $k=j$, $\tilde{\sigma}_j=(1-\alpha)/(\alpha.s_{ij}+(1-\alpha).s_{jj})$
Now define the 2x2 diagonal matrix $\tilde{\Sigma}_1=diag(\tilde{\sigma}_i,\tilde{\sigma}_j)$ and use row and column permutations upon the matrix $S=E.\Sigma^{-1}.E$ to move the elements $s_{ii},s_{ij}=s_{ji}, s_{jj}$ to the upper 2x2 part of the matrix. Let the result be the matrix:

$$\begin{bmatrix} S_{11} & S_{12} \\ t & \\ S_{12} & S_{22} \end{bmatrix}$$

where $S_{11} = \begin{bmatrix} s_{ii} & s_{ij} \\ s_{ji} & s_{jj} \end{bmatrix}$

From lemma 9 it follows that this matrix is PD.

Now $S_{11}^{-1} = 1/\det \begin{bmatrix} s_{jj} & -s_{ij} \\ -s_{ji} & s_{jj} \end{bmatrix}$ where $\det = s_{ii}.s_{jj} - s_{ij}^2$

and

$$S_{11}^{-1} - \tilde{\Sigma}_1 = s_{ij}/\det . \begin{bmatrix} x_j/x_i & -1 \\ -1 & x_i/x_j \end{bmatrix}$$

It is easy to see that $cor(S_{11}^{-1}-\tilde{\Sigma}_1)=1$ and that $(S_{11}^{-1}-\tilde{\Sigma}_1)$ is NND.
From lemma 11 we have : $cor(S_{11}^{-1}-\tilde{\Sigma}_1)=cor(\Sigma - \tilde{\Sigma})$
where

$$\tilde{\Sigma} = diag(0,0,\ldots,\tilde{\sigma}_i,0,\ldots,\tilde{\sigma}_j,0,\ldots 0)$$

From lemma 13 it follows that $\Sigma - \tilde{\Sigma}$ is NND.

Q.E.D.

In this section, we have studied the properties of vectors that conditions for such vectors to be NI, OI, allowed or not allowed. These properties will be used in a forthcoming paper to characterize the global solution set of the problem.

7. Conclusion.

The main goal of this paper was to introduce a novel geometrical approach for the maximal corank problem in the analysis of linear relations between noisy data.
New geometrical concepts, such as orthant invariance, null invariance were introduced. The general solution set consists of all vectors that are allowed, orthant null invariant.
We studied in detail the properties of vectors that are a convex combination of two least squares solutions and that have one zero component.
In a forthcoming publication, it will be demonstrated that these vectors are of key importance in the geometrical construction of the global solution set and that they can be used to prove Kalman's results [KAL] without using the Perron Frobeniustheorem.

8. References.

1.[DEM1] De Moor B., Vandewalle J. The uniqueness versus the non-uniqueness principle in the identification of linear relations from noisy data. Presented at this 25th IEEE CDC conference, Athens, dec.1986.
2.[DEM2] De Moor B., Vandewalle J. On the use of the inverse of the covariance matrix in multivariate data analysis. Submitted for publication, 1986.
3.[FRI] Frisch R. Statistical Confluence Analysis by means of complete regression systems. Publication nr.5, University of Oslo, Economic Institute 1934.
4.[GAN] Gantmacher F.R. The theory of matrices. Vol 1 & 2. Chelsea Publishing Company,New York.
5.[KAL] Kalman R.E. System identification from noisy data. Dynamical systems II. A.R. Bednarek and L. Cesari, editors. Academic Press, 1982.
6.[KOO] Koopmans T. Linear regression analysis of economic time series. Netherlands Economic Institute, Haarlem 1937.
7.[REI] Reiersol O. Confluence analysis by means of lag moments and other methods of confluence analysis. Econometrica 9, 1-24.