

A Rigorous Proof for the Invariant Study of the Problem in Section VI in “Matched Subspace Detectors”

Fangyuan Nan, Yaonan Wang and Fuhai Li
College of Electrical and Information Engineering,
Hunan University, Changsha, 410082, China

Abstract: Problem statement: This research provided rigorous proof for the invariant study of the problem in section VI in “matched subspace detectors”. About 14 years ago an important research entitled “matched subspace detectors” was published in IEEE transactions on signal processing, Vol. 42, No. 8, August 1994. Since its publication, the study has been widely cited in many areas. The main contribution of the research is to use invariance principle to study the Generalized Likelihood Ratio Test (GLRT) for four kinds of signal detection problems. While the conclusions are all correct, the largest invariant transformation group provided by the geometrical method is questionable. Furthermore the geometrical method in proving the maximal invariants is not helpful. The researchers themselves also frankly acknowledged “a rigorous proof requires an algebraic proof” (page 2152 in above research). **Approach:** Hence, this correspondence exactly gave rigorous proof based on algebraic method regarding one problem in the above mentioned research. **Results:** The algebraic method in this correspondence can be readily applied to other cases in the same research. **Conclusion/Recommendations:** Through this concrete example, we advocated the algebraic, rigorous method in the invariant study of signal detection problems, while abandoning the geometrical method.

Key words: Invariant test, invariant transformation group, maximal invariant, signal detection, matrix equation

INTRODUCTION

The invariance principle in dealing with hypothesis testing problem (or signal detecting problem) was described in several nice texts^[1-3]. Invariance study is a way impose constraints on detectors so that outputs are not dependent on convention choices such as units, coordinate systems and secondly so that the impact of nuisance parameters is minimized. The second is helpful for systematically obtaining potential properties such as constant false alarm rate.

Application of invariance principle to signal detection problem can be traced back as early as 1971^[4], or even earlier. About fourteen years ago an important paper^[5] appeared in IEEE Transaction on signal processing. Since its publication, the study been widely cited in many areas^[6-8]. The main contribution of the research is to use invariance principle to study the Generalized Likelihood Ratio Tests (GLRT) for four kinds signal detection problem. While the conclusion is all correct, the geometrical method used in this study is questionable as detailed in the next two paragraphs.

In invariance study of hypothesis testing problem, the first step is to find the invariant transformation group under which the problem is invariant and thus it is very important. In^[2,5], the transformation groups are given by some geometrical reasoning without solid proof. While the geometrical is illustrative of the invariant transformation group and invariant quantity (after they are found) and thus it is helpful for understanding and intuition, is defective in some aspects. It is not rigorous, as^[5] itself frankly acknowledged that “a rigorous proof requires an algebraic proof”. It is not general either. When applied to many complex, real-world problems, like in^[1], it is not productive.

Furthermore, the proof of maximal part of the invariance quantity is usually difficult and the geometrical method is not helpful either in this aspect. One can appreciate this point by reading^[1] to see how painstakingly the author was in proving the maximal part of the invariance quantity in many examples of that book. Because of these flaws, the geometrical method may mislead a novice just coming to this area.

Corresponding Author: Fangyuan Nan, College of Electrical and Information Engineering, Hunan University, Changsha, 410082, China

Hence, the following section in this paper exactly gives a rigorous proof based on algebraic method regarding the problem in section VI of^[5]. Another note should be mentioned here that^[3] implicitly finds subgroups of transformation and then builds up the larger transformation group by composition. Contrasted to^[3], present approach is a direct procedure based on solution to matrix equation. Contrasted to^[2,5], our method is more rigorous and more general.

MATERIALS AND METHODS

Problem statement: The detection problem in section VI of^[5] is briefly restated here, for sake of self-contained-ness. That is, given observation model:

$$Y^{N \times 1} = \mu X^{N \times 1} + S^{N \times t} \phi^{t \times 1} + n^{N \times 1} \tag{1}$$

where, n is Gaussian distributed: $n \sim N(0, \sigma^2 I)$.

We need decide between two hypotheses:

$$\begin{cases} H_0 : \mu = 0 \\ H_1 : \mu > 0 \end{cases} \tag{2}$$

Derivation of the invariant transformation group based on algebraic method: The following derivation and proofs all follow some relevant definition and theorems in chapter 6 of^[1]. In order to keep the family of probability density function (pdf) of y invariant, obviously we need choose a linear transformation on y , i.e.:

$$\gamma Q_M Y = \gamma \mu Q_M X + \gamma Q_M^S \phi + \gamma Q_M n \tag{3}$$

Where:

γ = A scalar

$Q_M^{N \times N}$ = A square matrix, which to be determined

In order to keep the structure of pdf of y invariant, obviously we need make sure that the followings hold:

$$\begin{cases} Q_M X = X \\ \gamma Q_M^S \phi = S \phi_1 \end{cases} \tag{4}$$

With $\phi_1 = D \phi$ and D is a matrix to be determined; in order to keep y still Gaussian distributed, obviously Q_M needs to be an orthogonal matrix.

Summing up, we need solve for an orthogonal matrix Q_M which satisfies:

$$\begin{cases} Q_M X = X \\ \gamma Q_M^S = S D \end{cases} \tag{5}$$

These two equations can be combined into one:

$$Q_M \begin{bmatrix} X & \gamma S \end{bmatrix} = \begin{bmatrix} X & S D \end{bmatrix} \tag{6}$$

At this stage we need exploit an important lemma.

Lemma 1 [1, Theorem A 9.5]: Suppose $A^{k \times m}$ and $B^{k \times n}$ are real matrices with $m \leq n$. A necessary and sufficient condition for $AA^T = BB^T$ is that there exists another quasi-orthogonal matrix $X^{m \times n}$, which means $XX^T = I^{m \times m}$, such that $AX = B$. IF $m = n$, then X is orthogonal.

A direct consequence of applying this lemma to Eq. (6) is $SD = \gamma S$ which simply implies $D = \gamma I$.

Under this condition, Eq. 6 is simplified to:

$$Q_M \begin{bmatrix} X & S \end{bmatrix} = \begin{bmatrix} X & S \end{bmatrix} \tag{7}$$

What is then the solution of Q_M to this equation in terms of x and S ? We need invoke another important lemma.

Lemma 2^[9]: Let A^+ and B^+ be pseudo-inverses of A and B , the linear matrix equation $AZB = E$, for matrix z is consistent if and only if $AA^+EB^+B = E$. Further, if the consistent condition is satisfied. The general solution is given by $Z = AA^+EB^+R^T - A^+AR^TB^+$ with R^T an arbitrary matrix.

It is an easy matter to check that the consistent condition for Eq. (7) to have a solution is always satisfied and therefore from Lemma 2 the general solution to Eq. (7) is:

$$Q_M = \begin{bmatrix} X & S \end{bmatrix} \begin{bmatrix} X & S \end{bmatrix}^+ R^T - R^T \begin{bmatrix} X & S \end{bmatrix} \begin{bmatrix} X & S \end{bmatrix}^+ \tag{8}$$

$$= P_M + R^T - R^T P_M = P_M + R^T P_M^\perp \tag{9}$$

where, $M = \begin{bmatrix} X & S \end{bmatrix}$, T means transpose, $^+$ means pseudo-inverse and P_M is the projection matrix as defined in^[2,5], that is $P_M = MM^+ = M(M^T M)^{-1} M^T$.

Do not forget we need make sure that Q_M is orthogonal. Let us first check $Q_M Q_M^T = I$, which leads to:

$$R^T P_M^\perp R = P_M^\perp \tag{10}$$

We decompose:

$$\begin{cases} P_M = U_{M2}^{N \times (t+1)} U_{M2}^T \\ P_M^\perp = U_{M1}^{N \times (N-t-1)} U_{M1}^T \end{cases} \tag{11}$$

where, U_{M2} is a quasi- orthogonal matrix which means $U_{M2}^T U_{M2} = I$. So is the case for U_{M1} . These decompositions follow from page 48 in^[2]. Furthermore we define:

$$U \square [U_{M1} \quad U_{M2}]$$

which is an orthogonal matrix.

Based on above decomposition, Eq. 10 becomes:

$$R^T U_{M1} U_{M1}^T R = R^T U_{M1} (R^T U_{M1})^T = U_{M1} U_{M1}^T \quad (12)$$

Therefore, from Lemma 1, there exists an orthogonal matrix Q_1 such that:

$$R^T = U_{M1} U_{M1} Q_1 \quad (13)$$

which implies $R^T P_M^\perp = R^T U_{M1} U_{M1}^T = U_{M1} Q_1 U_{M1}^T$.

Substitution of this result into Eq. (9) yields:

$$Q_M = P_M + R^T P_M^\perp = U_{M1} Q_1 U_{M1}^T + U_{M2} U_{M2}^T \quad (14)$$

$$= [U_{M1} \quad U_{M2}] \begin{bmatrix} Q_1 & 0 \\ 0 & I^{T \times T} \end{bmatrix} \begin{bmatrix} U_{M1}^T \\ U_{M2}^T \end{bmatrix} \quad (15)$$

Q_M is obviously an orthogonal matrix, actually a rotation matrix^[2] in subspace $\langle M \rangle$, where, following the notation in^[5], $\langle M \rangle$ denotes the subspace spanned by columns of M . This justifies our notation which is in agreement with that used in^[2].

To match the result in^[5], we need show that Q_M is the rotation matrix in subspace $\langle P_S^\perp x \rangle^\perp$. we first define $G^{N \times 1} = P_S^\perp x$, correspondingly $U_G = G/\|G\|$ and define another matrix $F^{N \times (N-1)}$ such that $\langle F \rangle = \langle P_S^\perp x \rangle^\perp$. As usual, we decompose $P_F = U_F U_F^T$. According to Eq. (3.2) in^[5]:

$$P_M = P_S + P_G = U_G U_G^T + U_M U_M^T \quad (16)$$

The beauty of this decomposition is that $P_S \perp P_G$, explicitly written as $P_S^T P_G = 0$.

If the order of columns in U_M is not taken into account, this decomposition is unique^[2]. So one column in U_M , say, the last one, is U_G and all the other columns in U_M correspond to those in U_S .

Therefore:

$$Q_M = [U_{M1} \quad U_{M2}] \begin{bmatrix} Q_1 & 0 & 0 \\ 0 & I^{T \times T} & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} U_{M1}^T \\ U_{M2}^T \end{bmatrix} \quad (17)$$

$$Q_M = [U_F \quad U_G] \begin{bmatrix} Q_2 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} U_F^T \\ U_G^T \end{bmatrix} \quad (18)$$

The last equality results from the following reasoning. Denote all columns in U as U_{M3} except for the last one which was already shown as the same as U_G . Therefore:

$$I = U_{M3} U_{M3}^T + U_G U_G^T = U_F U_F^T + U_G U_G^T \quad (19)$$

This is tantamount to saying that:

$$U_{M3} U_{M3}^T = U_F U_F^T = I - U_G U_G^T = P_G^\perp \quad (20)$$

Because P_G^\perp is fixed, this decomposition is unique, that is, $U_{M3} = U_F$.

And Eq. 18 is exactly the rotation matrix in subspace $\langle P_S^\perp x \rangle^\perp = \langle F \rangle$. For the definition of a rotation matrix in subspace, please refer to^[2].

So we proved rigorously (algebraically) the conclusion in^[5] that the detection problem is invariant to set of rotation in $\langle P_S^\perp x \rangle^\perp$ and scaling $\gamma > 0$.

It is not hard to check that the pdf of y remains invariant under the group of transformation of γQ_M .

RESULTS AND DISCUSSION

Algebraic proof of the invariance of $L_2(Y)$ in section VI of^[5]: It is claimed in^[5] that, for the detection problem considered in this correspondence, a maximal invariant is:

$$L_2(y) = \frac{y^T P_S^\perp P_G P_S^\perp y}{y^T P_S^\perp P_G^\perp P_S^\perp y} \quad (21)$$

However, Scharf and Friedlander^[5] did not give any proof and, as stated in the introduction of this correspondence, the geometrical method used in is not helpful in proving the maximal invariant.

We will see that the algebraic method employed in this study is fruitful in this regard. We first check on the invariance. The transformation by a constant scaling $\gamma > 0$ obviously leaves $L_2(y)$ changed. Therefore we consider

$$L_2(Q_M y) = \frac{y^T (U_{M2} U_{M2}^T + U_{M1} Q_1^T U_{M1}^T) U_G U_G^T (U_{M2} U_{M2}^T + U_{M1} Q_1 U_{M1}^T) y}{y^T (U_{M2} U_{M2}^T + U_{M1} Q_1^T U_{M1}^T) U_{M1} U_{M1}^T (U_{M2} U_{M2}^T + U_{M1} Q_1 U_{M1}^T) y} \quad (22)$$

Paying attention to the fact that $U_{M_2} \perp U_{M_1}$ and that U_G is the last column in U_M , we can see the above equation is simplified to:

$$L_2(Q_{MY}) = \frac{y^T (U_{M_2} U_{M_2}^T U_G) (U_G^T U_{M_2} U_{M_2}^T) y}{y^T (U_{M_1} Q_1^T I I Q_1 U_{M_1}^T) y} \quad (23)$$

$$= \frac{y^T (U_{M_2} U_{M_2}^T U_G) (U_G^T U_G) (U_G^T U_{M_2} U_{M_2}^T) y}{y^T U_{M_1} U_{M_1}^T y} \quad (24)$$

$$= \frac{(Y^T P_M P_G) (P_G^T P_M^T) Y}{y^T P_{[X \ s]}^{\perp} y} \quad (25)$$

Because $P_M P_G = (P_S + P_G) P_G = 0 + P_G$:

$$L_2(Q_{MY}) = \frac{Y^T P_G P_G^T Y}{y^T P_{[X \ s]}^{\perp} y} \quad (26)$$

$$= \frac{Y^T P_S^{\perp} P_G P_S^{\perp} Y}{y^T P_S^{\perp} P_G P_S^{\perp} y} = L_2(Y) \quad (27)$$

The second equality follows from Eq. 3.6 in^[5]. Therefore we complete the proof that $L_2(Y)$ is invariant under the group of transformation prescribed in the previous subsection.

Algebraic proof of the maximal invariant part of $L_2(Y)$ in section VI of^[5]: As stated in the introduction of this correspondence, the geometrical method used in^[5] is helpless in proving the maximal invariant. Here we give a rigorous proof based on algebraic method. Suppose we are given Y_2, Y_1 such that:

$$L_2(Y_2) = L_2(Y_1) \quad (28)$$

We need show there exists a constant γ and an orthogonal matrix Q_1 such that:

$$Y_2 = \gamma U \begin{bmatrix} I & \\ & Q_1 \end{bmatrix} U_{Y_1}^T \quad (29)$$

Obviously we need pick up:

$$\gamma = \frac{\|y_2\|}{\|y_1\|} \quad (30)$$

Expanding Eq. 29 leads to:

$$\frac{y_2}{\|y_2\|} = (I - U_{M_1} U_{M_1}^T + U_{M_1} Q_1 U_{M_1}^T) \frac{y_1}{\|y_1\|} \quad (31)$$

Which is equivalent to:

$$\frac{y_2}{\|y_2\|} - (I - U_{M_1} U_{M_1}^T) \frac{y_1}{\|y_1\|} = U_{M_1} Q_1 U_{M_1}^T \frac{y_1}{\|y_1\|} \quad (32)$$

Left multiplying the above equation by $U_{M_1}^T$ result in:

$$\frac{y_1^T}{\|y_1\|} U_{M_1} Q_1^T = \frac{y_2^T}{\|y_2\|} U_{M_1} \quad (33)$$

By lemma 1, a necessary and sufficient condition for the above equation to have an orthogonal matrix solution of Q_1^T is:

$$\frac{y_1^T}{\|y_1\|} U_{M_1} U_{M_1}^T \frac{y_1}{\|y_1\|} = \frac{y_2^T}{\|y_2\|} U_{M_1} U_{M_1}^T \frac{y_2}{\|y_2\|} \quad (34)$$

By some algebra, the above equation is equivalent to:

$$\frac{\|P_{S Y_1}\|^2}{\|P_S^{\perp} P_G^{\perp} P_S^{\perp} y_1\|^2} = \frac{\|P_{S Y_2}\|^2}{\|P_S^{\perp} P_G^{\perp} P_S^{\perp} y_2\|^2} \quad (35)$$

From Eq. 28:

$$= \frac{y_1^T P_{G Y_1}}{y_1^T P_S^{\perp} P_G^{\perp} P_S^{\perp} y_1} = \frac{y_2^T P_{G Y_2}}{y_2^T P_S^{\perp} P_G^{\perp} P_S^{\perp} y_2} \quad (36)$$

Adding 1 to both sides of Eq. 36:

$$\frac{y_1^T P_S^{\perp} y_1}{y_1^T P_S^{\perp} P_G^{\perp} P_S^{\perp} y_1} = \frac{y_2^T P_{G Y_2}}{y_2^T P_S^{\perp} P_G^{\perp} P_S^{\perp} y_2} \quad (37)$$

which can be shown to lead to Eq. 29.

CONCLUSION

In this short correspondence, we consider one signal detection problem in^[5] by invariance principle. Contrasted to the geometrical method in^[5], our method is rigorous, algebraic (solution to matrix equation are often used), thereby completing^[5]. Our method is also more general, because the same methodology can be readily applied to other cases in^[2,5].

REFERENCES

1. Muirhead, R.J., 1982. Aspects of Multivariate Statistical Theory. 1st Edn., John Wiley and Sons, USA., ISBN: 10: 0471094420, pp: 704.

2. Scharf, L.L., 1990. *Statistical Signal Processing, Detection, Estimation and Time Series Analysis*. 1st Edn., Addison Wesley, ISBN: 10: 0201190389, pp: 524.
3. Lehman, E.L. and J.P. Romano, 2005. *Testing Statistical Hypothesis*. 3rd Edn., Springer, New York, pp: 300-350.
4. Scharf, L.L. and D.W. Lytle, 1971. Signal detection in Gaussian noise of unknown level: An invariance application. *IEEE Trans. Inform. Theor.*, IT: 404-411. http://ieeexplore.ieee.org/xpl/freeabs_all.jsp?arnumber=1054668
5. Scharf, L.L. and B. Friedlander, 1994. Matched subspace detectors. *IEEE Trans. Signal Process.*, 42: 2146-2157. DOI: 10.1109/78.301849
6. Nan, F.Y. and R.D. Nowak, 1999. Generalized likelihood ratio detection for fMRI using complex data. *IEEE Trans. Med. Imag.*, 18: 320-329. DOI: 10.1109/42.768841
7. Conte, E., A.D. Maio and C. Galdi, 2003. CFAR detection of multidimensional signal: An invariant approach. *IEEE Trans. Signal Process.*, 51: 142-151. DOI: 10.1109/TSP.2002.806554
8. Row, D.B. and B.R. Logan, 2004. A complex way to compute fMRI activation. *Neuroimage*, 23: 1078-1092. <http://www.ncbi.nlm.nih.gov/pubmed/15528108>
9. Shim, S.Y. and Y. Chen, 2002. Least squares solution of matrix equation $AXB^* + CYD^* = E^{**}$. *SIAM J. Matrix Anal.*, 24: 802-808. <http://portal.acm.org/citation.cfm?id=636344>