Use of Wijsman's Theorem for the Ratio of Maximal Invariant Densities in Signal Detection Applications

Joseph R. Gabriel
Naval Undersea Warfare Center
Newport, RI 02841

Steven M. Kay
University of Rhode Island
Kingston, RI 02881

Abstract

We apply Wijsman's theorem for the ratio of densities of the maximal invariant to signal detection applications. The method does not require the explicit use of a maximal invariant statistic or its density to derive the uniformly most powerful invariant (UMPI) test. We describe its use for common classes of detection problems and illustrate its use in examples. The analytic form of the representation provides new insight into the relationship between the UMPI and generalized likelihood ratio tests.

1 Introduction

1.1 Hypothesis testing

The criterion used to select a test for signal detection applications such as sonar and radar is typically the test that provides the highest probability of detection (Pd) subject to a given probability of false alarm (Pfa) requirement. The detection problem is cast as a hypothesis test in which the null hypothesis is noise and the alternate hypothesis is signal plus noise. However, use of the Neyman-Pearson lemma to derive tests for these applications usually does not result in a test that can be implemented due to unknown signal or noise parameters. A uniformly most powerful (UMP) test, in which the test has the highest Pd over the domain of unknown parameters, also usually does not exist.

When the UMP test does not exist, various approaches can be used to select a suboptimal test. One of the most widely used is the generalized likelihood ratio test (GLRT). To derive the GLRT, the unknown parameters under each hypothesis are replaced by their maximum likelihood estimates (MLE's) [1]. Among the known properties of the GLRT in invariant problems are that it is invariant and asymptotically UMPI. However, it is not necessarily UMPI for finite data records, even when the UMPI test exists.

1.2 Invariance

The principle of statistical invariance is applied to composite hypothesis testing problems, and in particular those for which a UMP test does not exist, as an approach to find the most powerful test among the class of invariant tests. Using the criterion that the test be invariant, as a further requirement of a satisfactory test, may be reasonable for the problem and lead to a test with optimality properties. Even if the UMPI test is found not to exist, then the derived result can be used as a performance bound for the class of all invariant tests.

Application of invariance principles to hypothesis testing is described in statistics texts such as Lehmann [2]. The text of Scharf [3] applies invariance methods to signal detection applications. The classic approach of finding the UMPI test is to: identify problem invariances, describe the invariance analytically in terms of a transformation group, choose a maximal invariant statistic, derive the probability density function (PDF) of the maximal invariant under each hypothesis, and then construct the likelihood ratio. Another method of deriving the UMPI test statistic, is to show that a scalar maximal invariant has a monotone likelihood ratio. Scharf and Friedlander [4] use that approach to show that the GLRT is UMPI for classes of detection problems that fit the linear model. Here, we use Wijsman's theorem for the likelihood ratio of maximal invariant densities, which does not require a maximal invariant or its density. We note that the analytic form of the ratio provides insight into the relationship between the UMPI and suboptimal tests. In some problems, the UMPI test statistic can be interpreted to be an average of the likelihood ratio over the group and the GLRT statistic the maximum of the ratio over the group.

0-7803-7576-9/02$17.00

1.3 Principal literature

Stein [5] is credited in the statistics literature for the general method of averaging over a group to obtain the density of a maximal invariant statistic. He also provided an expression for the ratio of maximal invariant densities. Wijsman [6] further developed the approach using a theory of cross sections of orbits and providing conditions under which the ratio is valid. In [7] Wijsman gives alternate conditions which can be easier to apply using the concept of proper actions (Andersson [8]) of a group. Using numerous examples, he also shows that the conditions are met for all the "usual" problems arising in normal multivariate analysis. Eaton [9] provides a description and examples of the approach, including a version of Wijsman's theorem. We apply these results to the groups used in signal processing applications, show that the ratio validity conditions are met, and illustrate its use for common classes of signal detection problems.

Applications of the approach to obtain the density or ratio of maximal invariants can be found in the statistics literature (e.g., in Uthoff [10] and Kariya and Sinha [11]). The only use of the approach in the engineering literature known by us is Schwartz [12]. He uses invariance principles to derive CFAR and minimax detectors for a signal of unknown phase in Gaussian noise of unknown covariance. To obtain the ratio of maximal invariant densities, he integrates over a group (which is generated by the composition of two groups). The integral form is then approximated for the low SNR case to obtain a CFAR locally minimax test.

2 Representation of the ratio of maximal invariant densities

To derive an UMPI test, the problem invariances and a group G under which transformations on the data should leave the problem invariant are identified. The group elements are denoted as $g \in G$. We next use Wijsman's theorem to derive the ratio of maximal invariant densities. In particular, we use the approach as described in Eaton [9] (theorems 5.8 and 5.9) in which the ratio is given by

$$\frac{\int p_{\mathbf{x}}\left(g\mathbf{x};\mathcal{H}_1\right)\chi_0\left(g\right)d\nu_l\left(g\right)}{\int p_{\mathbf{x}}\left(g\mathbf{x};\mathcal{H}_0\right)\chi_0\left(g\right)d\nu_l\left(g\right)}.$$

Lower case bold notation is used to denote vectors and later, upper case bold will be used to denote matrices. The notation $g\mathbf{x}$ indicates group element transformations on the data, and $\bar{g}\boldsymbol{\theta}$ is the notation used to denote induced group actions on the parameter space. To interpret this ratio, first consider that given that the problem is invariant under the transformation group G, then densities under each hypothesis must satisfy the relationship $p_{\mathbf{x}}\left(g^{-1}\mathbf{x};\boldsymbol{\theta}\right)\left|J_{g^{-1}}\left(\mathbf{x}\right)\right| = p_{\mathbf{x}}\left(\mathbf{x};\bar{g}\boldsymbol{\theta}\right)$. The Jacobian $\left|J_{g^{-1}}\left(\mathbf{x}\right)\right|$ is the usual factor obtained in transformation of random variables. The factor $\chi_0\left(g\right)$ of Wijsman's ratio is the inverse of this Jacobian, $\left|J_{g^{-1}}\left(\mathbf{x}\right)\right|^{-1}$. The measure $\nu_l\left(g\right)$ is a left invariant measure for the group. For signal processing applications, this measure can usually be written in terms of a Lebesgue measure, making it possible to evaluate the resulting integral form. For cases of discrete groups, the integral is easily converted into a summation. See Eaton [9] and Kariya [11] for other examples.

2.1 Conditions

Wijsman [7] describes the conditions under which the ratio is valid to be that the group is a Lie group, the sample space is a subset of Euclidean space, and the action is linear or affine. He states that these are not much of a limitation in practice since all known examples seem to be of this nature. Additionally, the group action must be proper, and under the assumption of locally compact groups he provides a lemma that is easy to apply. Eaton provides a version of this lemma as Theorem 5.6. Wijsman also shows how to apply his lemma to groups that have been generated by subgroups, which is analogous to a theorem in Lehmann (Chap 6, Thm 2) in which a maximal invariant is generated in steps. This is important since we are now working with groups instead of maximal invariant statistics.

Appendix I provides descriptions and examples for the lemma terms of proper action and locally compact. A summary of common groups and those relevant to signal processing applications is also provided. All of the groups shown meet the theorem conditions. In using this approach to derive UMPI tests, we work with groups, subgroups, and group compositions, instead of with maximal invariant statistics and their densities.

3 Signal detection examples

Two examples are provided in which the UMPI test is derived and compared with the GLRT. Insights gained by interpretation of the analytic form of the integral are provided.

Example 1 *Signal Known Except for Amplitude*

The hypothesis test is

$$\begin{aligned} \mathcal{H}_0 &: \mathbf{x} = \mathbf{w} \\ \mathcal{H}_1 &: \mathbf{x} = A\mathbf{s} + \mathbf{w}, \end{aligned}$$

where $\mathbf{s}$ is a known signal, $\mathbf{w}$ is the noise vector which is distributed as $\mathcal{N}(\mathbf{0}, \sigma^2\mathbf{I})$, and A $(-\infty < A < \infty)$ is an unknown scaling factor on the signal. The noise variance σ^2 is known. $\mathcal{H}_0$ and $\mathcal{H}_1$ denote the null and alternate hypotheses respectively. The UMP test for this problem is $\operatorname{sgn}(A)\mathbf{s}^T\mathbf{x} > \gamma$, where γ is the threshold established by the Pfa requirement. Hence, for the UMP test to exist, we need to know $\operatorname{sgn}(A)$.

Since we do not know $\operatorname{sgn}(A)$, then we should consider tests invariant to the sign. The discrete group $G_s = \{g_s : g_s = -1, 1\}$ acts on the data samples $\mathbf{x}$ under the operation of multiplication. To establish that the problem is invariant [2], it is shown in Appendix II that the densities belong to the same family after transformation, and the parameter spaces are preserved.

As noted in section II.B and in Appendix I, this group is locally compact and acts properly since it is discrete. Hence the conditions of Wijsman's theorem are met. For this group, $\chi_0(g_s) = 1$ since $p_{\mathbf{x}}(g_s\mathbf{x}; \bar{g}_s\boldsymbol{\theta}) = p_{\mathbf{x}}(\mathbf{x}; \boldsymbol{\theta})$. A counting measure is a left invariant measure and used since the group is discrete. Hence, the ratio of maximal invariant densities is given by $\frac{\sum_{g_s=\{-1,1\}} p_{\mathbf{x}}(g_s\mathbf{x};\mathcal{H}_1)}{\sum_{g_s=\{-1,1\}} p_{\mathbf{x}}(g_s\mathbf{x};\mathcal{H}_0)} = \frac{1}{2}\sum_{g_s=\{-1,1\}} \frac{p_{\mathbf{x}}(g_s\mathbf{x};\mathcal{H}_1)}{p_{\mathbf{x}}(\mathbf{x};\mathcal{H}_0)}$, since $p_{\mathbf{x}}(\mathbf{x};\mathcal{H}_0) = p_{\mathbf{x}}(-\mathbf{x};\mathcal{H}_0)$. Expanding the expression gives

$$\begin{aligned} &\frac{1}{2}\sum_{g_s=\{-1,1\}} \frac{\exp\left[-\frac{1}{2\sigma^2}(g_s\mathbf{x} - A\mathbf{s})^T(g_s\mathbf{x} - A\mathbf{s})\right]}{\exp\left[-\frac{1}{2\sigma^2}\mathbf{x}^T\mathbf{x}\right]} \\ &= \frac{1}{2}\exp\left[-\frac{A^2}{2\sigma^2}\mathbf{s}^T\mathbf{s}\right]\left(\exp\left[\frac{A}{\sigma^2}\mathbf{s}^T\mathbf{x}\right] + \exp\left[\frac{-A}{\sigma^2}\mathbf{s}^T\mathbf{x}\right]\right). \end{aligned}$$

The expression is monotonic with respect to $|\mathbf{s}^T\mathbf{x}|$, which can be used as the UMPI test statistic. Observe that the above expression is a summation over a two element group which can be interpreted as an average over the group. The GLRT will be seen to have an interpretation of the maximum over the group.

The GLRT is obtained by substituting the MLE of A, which is $\hat{A} = (\mathbf{s}^T\mathbf{s})^{-1}\mathbf{s}^T\mathbf{x}$, into the likelihood ratio,

$$\begin{aligned} &\frac{p_{\mathbf{x}}\left(\mathbf{x}; \hat{A}\mathbf{s}, \mathcal{H}_1\right)}{p_{\mathbf{x}}(\mathbf{x};\mathcal{H}_0)} \\ &= \exp\left[\frac{-1}{2\sigma^2}\left(\left(g_s\mathbf{x} - \hat{A}\mathbf{s}\right)^T\left(g_s\mathbf{x} - \hat{A}\mathbf{s}\right) - \mathbf{x}^T\mathbf{x}\right)\right] \\ &= \exp\left[\frac{1}{2\sigma^2}\left(\mathbf{s}^T\mathbf{s}\right)^{-1}\left(\mathbf{s}^T\mathbf{x}\right)^2\right]. \end{aligned}$$

This is also monotonic with respect to $|\mathbf{s}^T\mathbf{x}|$, and hence the GLRT and the UMPI tests are the same.

Next consider the case in which it is given that $A > 0$. Here, the UMP test exists and is $\mathbf{s}^T\mathbf{x}$ since we know $\operatorname{sgn}(A)$. The two element group G_s used in the $A \neq 0$ case reduces to a group with a single element, $G = \{1\}$. This trivial case is shown for illustration of the concept, and will be extended later. Using this group in the ratio gives $\frac{\sum_{g=\{1\}} p_{\mathbf{x}}(g\mathbf{x};\mathcal{H}_1)}{\sum_{g=\{1\}} p_{\mathbf{x}}(g\mathbf{x};\mathcal{H}_0)} = \frac{\exp\left[\frac{-1}{2\sigma^2}(\mathbf{x}-A\mathbf{s})^T(\mathbf{x}-A\mathbf{s})\right]}{\exp\left[\frac{-1}{2\sigma^2}\mathbf{x}^T\mathbf{x}\right]} = \exp\left[\frac{A}{\sigma^2}\mathbf{s}^T\mathbf{x} - \frac{A^2}{2\sigma^2}\mathbf{s}^T\mathbf{s}\right]$, which is increasing in $\mathbf{s}^T\mathbf{x}$, the UMPI (and UMP) test statistic. However, the GLRT for this example is not UMPI. The MLE for A is $\hat{A} = \max\left(0, (\mathbf{s}^T\mathbf{s})^{-1}\mathbf{s}^T\mathbf{x}\right)$. The likelihood ratio is

$$\begin{aligned} \frac{p_{\mathbf{x}}\left(\mathbf{x}; \hat{A}\mathbf{s}, \mathcal{H}_1\right)}{p_{\mathbf{x}}(\mathbf{x};\mathcal{H}_0)} &= \exp\left[\frac{-1}{2\sigma^2}\left(-2\hat{A}\mathbf{s}^T\mathbf{x} + \hat{A}^2\mathbf{s}^T\mathbf{s}\right)\right] \\ &= \begin{cases} \exp[0], \text{ for } \hat{A} = 0 \\ \exp\left[\frac{1}{2\sigma^2}\left(\mathbf{s}^T\mathbf{s}\right)^{-1}\left(\mathbf{s}^T\mathbf{x}\right)^2\right], \text{ for } \hat{A} > 0 \end{cases}. \end{aligned}$$

Taking the natural log, we see that the GLRT can be written as $\max(0, \mathbf{s}^T\mathbf{x})$. This test has a lower probability of detection (Pd) than the UMPI test for probability of false alarms (Pfa) in the interval $[.5, 1)$. Notice that in the GLRT derivation, the parameter A is replaced. The UMPI test obtained by integrating over the sign group, induces action only on the parameter $\operatorname{sgn}(A)$ and not A, which is $\operatorname{sgn}(A)|A|$. In replacing all parameters, the GLRT may imply a larger group invariance than indicated by the problem itself, and can result in a suboptimal test. Eaton ([9], example 6.5) provides a different example in which the GLRT is not UMPI and also attributes the result to a bit "too much invariance," noting that *fully invariant procedures such as the likelihood ratio test (GLRT) can be improved upon by simply requiring less invariance.*

We now extend the example beyond the trivial group to the case in which we have interference of unknown level. The hypothesis test is

$$\begin{aligned}\mathcal{H}_0 &: \mathbf{x} = d\mathbf{1} + \mathbf{w} \\ \mathcal{H}_1 &: \mathbf{x} = d\mathbf{1} + A\mathbf{s} + \mathbf{w},\end{aligned}$$

where the noise $\mathbf{w}$ is Gaussian with known variance $\sigma^2\mathbf{I}$, and the $\mathbf{1}$ denotes an $N \times 1$ vector whose elements are all 1. It is also given that $A > 0$, d is unknown, and $\mathbf{s} = \begin{bmatrix} 1 & -1 \end{bmatrix}^T$. The two sample case is considered for simplicity. The density is $\mathcal{N}(d\mathbf{1}, \sigma^2\mathbf{I})$ under $\mathcal{H}_0$, and $\mathcal{N}(d\mathbf{1}+A\mathbf{s}, \sigma^2\mathbf{I})$ under $\mathcal{H}_1$. The problem is invariant under the translation group, $G_b = \{g_b : g_b\mathbf{x} = \mathbf{x} + b\mathbf{1},\ b \in R^1\}$. The group transformations of $\mathbf{x}$ are Gaussian, with means ($b\mathbf{1}+d\mathbf{1}$ and $b\mathbf{1}+d\mathbf{1}+A\mathbf{s}$) and covariance matrices the same as before. The conditions of Wijsman's theorem are met since the group is locally compact and acts properly (Appendix I). The multiplier $\chi_0(g_b)$ is 1 since $p_\mathbf{x}(g_b\mathbf{x}; \bar{g}_b\boldsymbol{\theta}) = p_\mathbf{x}(\mathbf{x} + b\mathbf{1}; (d-b)\mathbf{1} + A\mathbf{s}) = p_\mathbf{x}(\mathbf{x}; d\mathbf{1}+A\mathbf{s}) = p_\mathbf{x}(\mathbf{x}; \boldsymbol{\theta})$. That is, the random variable transformation by the group elements has a Jacobian of 1. For the left invariant measure we use Lebesgue measure db since subintervals of equal length have equal measure along R^1 (see Appendix III for a different example). Transformation by translation group elements does not change the size of the subinterval and hence the measure is unchanged. Furthermore, since the integrand is continuous, the Lebesgue integral can be interpreted as a Riemann integral.

Using these results, the ratio is $\frac{\int_{-\infty}^{\infty} p_\mathbf{x}(\mathbf{x}+b\mathbf{1};\mathcal{H}_1)db}{\int_{-\infty}^{\infty} p_\mathbf{x}(\mathbf{x}+b\mathbf{1};\mathcal{H}_0)db}$, which is

$$\frac{\int_{-\infty}^{\infty} \exp\left[\begin{array}{c}\frac{-1}{2\sigma^2}(\mathbf{x}+(b-d)\mathbf{1}-A\mathbf{s})^T \\ \cdot(\mathbf{x}+(b-d)\mathbf{1}-A\mathbf{s})\end{array}\right] db}{\int_{-\infty}^{\infty} \exp\left[\frac{-1}{2\sigma^2}(\mathbf{x}+b\mathbf{1})^T(\mathbf{x}+b\mathbf{1})\right] db},$$

and simplified further to

$$= \frac{\int_{-\infty}^{\infty} \exp\left[\frac{-1}{2\sigma^2}(\mathbf{x}+b\mathbf{1}-A\mathbf{s})^T(\mathbf{x}+b\mathbf{1}-A\mathbf{s})\right] db}{\int_{-\infty}^{\infty} \exp\left[\frac{-1}{2\sigma^2}(\mathbf{x}+b\mathbf{1})^T(\mathbf{x}+b\mathbf{1})\right] db},$$

where $b\mathbf{1} - d\mathbf{1}$ has been replaced with $b\mathbf{1}$ since the integration is over the interval $(-\infty, \infty)$. The $A\mathbf{s}$ factor can be brought outside the integral since $\mathbf{s}^T\mathbf{1} = 0$ and it is not a function of b, giving

$$\exp\left[\frac{A}{\sigma^2}\mathbf{s}^T\mathbf{x} - \frac{A^2}{\sigma^2}\right] \frac{\int_{-\infty}^{\infty} \exp\left[\frac{-1}{2\sigma^2}(\mathbf{x}+b\mathbf{1})^T(\mathbf{x}+b\mathbf{1})\right] db}{\int_{-\infty}^{\infty} \exp\left[\frac{-1}{2\sigma^2}(\mathbf{x}+b\mathbf{1})^T(\mathbf{x}+b\mathbf{1})\right] db}.$$

This is increasing in $\mathbf{s}^T\mathbf{x}$, since $A > 0$. Hence, the UMPI test statistic is $\mathbf{s}^T\mathbf{x}$. Note that since $\mathbf{s} = \begin{bmatrix} 1 & -1 \end{bmatrix}^T$, the statistic $\mathbf{s}^T\mathbf{x} = x[1] - x[2]$, which is invariant to the interference.

The GLRT is obtained by substituting $\hat{A}$, $\hat{d}_0$, and $\hat{d}_1$, the MLEs of the parameters under each hypothesis into $\frac{p_\mathbf{x}(\mathbf{x};\mathcal{H}_1)}{p_\mathbf{x}(\mathbf{x};\mathcal{H}_0)}$. The MLE's are $\hat{A} = \max\left(0, \frac{1}{2}\mathbf{s}^T\mathbf{x}\right)$, and $\hat{d}_0 = \hat{d}_1 = \frac{1}{2}\mathbf{1}^T\mathbf{x}$. Making the substitutions and some cancellations gives $\exp\left[\frac{1}{4\sigma^2}\left(\mathbf{s}^T\mathbf{x}\right)^2\right]$ for the $\hat{A} > 0$ case, and 0 for the $\hat{A} = 0$ case. This is increasing in $\max(0, \mathbf{s}^T\mathbf{x})$, which can be used as the GLRT statistic.

Hence, this is an example for which *the UMPI test exists, but is not given by the GLRT.* As noted for the two-sided case, in deriving the GLRT we replaced all unknown parameters. In the UMPI derivation, integration over the translation group implies an induced action on the interference parameter only.

Example 2 *Signal in a Known Subspace*

This example is of the form of the classic linear model

$$\begin{aligned}\mathcal{H}_0 &: \mathbf{x} = \mathbf{w} \\ \mathcal{H}_1 &: \mathbf{x} = \mathbf{H}\boldsymbol{\theta} + \mathbf{w}.\end{aligned}$$

The $N \times p$ matrix $\mathbf{H}$ is known, the noise $\mathbf{w}$ is Gaussian with known variance σ^2, and the $p \times 1$ parameter vector $\boldsymbol{\theta}$ is unknown.

We use the $p \times 1$ sufficient statistic $\mathbf{z} = \left(\mathbf{H}^T\mathbf{H}\right)^{-\frac{1}{2}}\mathbf{H}^T\mathbf{x}$, and write the equivalent hypothesis test using $\mathbf{z} = \mathbf{w}$ under $\mathcal{H}_0$, and $\mathbf{z} = \left(\mathbf{H}^T\mathbf{H}\right)^{\frac{1}{2}}\boldsymbol{\theta} + \left(\mathbf{H}^T\mathbf{H}\right)^{-\frac{1}{2}}\mathbf{H}^T\mathbf{w}$ under $\mathcal{H}_1$. Since the noise term still has a density $N\left(\mathbf{0}, \sigma^2\mathbf{I}\right)$, the density of $\mathbf{z}$ is $\mathcal{N}\left(\left(\mathbf{H}^T\mathbf{H}\right)^{\frac{1}{2}}\boldsymbol{\theta}, \sigma^2\mathbf{I}\right)$ under $\mathcal{H}_1$ and $\mathcal{N}\left(\mathbf{0}, \sigma^2\mathbf{I}\right)$ under $\mathcal{H}_0$. Note that we use this version of the sufficient statistic since it is invariant under the group $G_o = \{g_o : g_o\mathbf{z} = \mathcal{O}\mathbf{z},\ \mathcal{O}^T\mathcal{O} = \mathbf{I}\}$. The matrix $\mathcal{O}$ is the $p \times p$ orthogonal group. The problem is shown to be invariant in Appendix II. Other statistics of the data and different groups can be used for this problem ([3] and [4]). The statistic and group used rotates the component of the data that lies in the signal subspace and leave other components unchanged.

The $p = 2$ case is now developed. Extension to the general p case can be easily made using generalized spherical coordinates. For this $p = 2$ case, the orthogonal group is a rotation in R^2 and can be written as $G_o = \left\{g_\eta : g_\eta\mathbf{z} = \begin{bmatrix} \cos\eta & -\sin\eta \\ \sin\eta & \cos\eta \end{bmatrix}\mathbf{z},\ \eta \epsilon [-\pi, \pi]\right\}$, which is a set of 2×2 matrices that act under multiplication on $\mathbf{z}$. Alternatively, consider the group $G_\gamma = \{g_\gamma : g_\gamma \mathbf{z}(r,\alpha) = \mathbf{z}(r,\gamma+\alpha), \gamma \epsilon [-\pi,\pi]\}$, which is a set of translation elements in R^1 acting under modulo addition on the angle elements (of the polar form) of $\mathbf{z}$. Define $\mathbf{z}(r,\alpha) = \begin{bmatrix} r\cos\gamma & r\sin\gamma \end{bmatrix}^T$. Since it can be shown that G_o and G_γ are isomorphic (i.e., there is a 1-1 correspondence between elements and also between the results of computations), we can use either group and its associated action on the data to describe the problem invariance. This is a translation group and hence is locally compact and acts properly (Appendix I). For the group G_γ, the left invariant measure $d\nu_l(g_\gamma)$ is the Lebesgue measure $d\gamma$ since the composite group actions are translations in R^1 and Lebesgue measure for an interval is invariant under translation. The group multiplier $\chi_0(g_\gamma)$ is 1.

The group acts on the angle component of $\mathbf{z}(r,\alpha)$. Under $\mathcal{H}_1$,

$$p_{\mathbf{z}}\left(g_\gamma \mathbf{z}(r,\alpha); (\mathbf{H}^T\mathbf{H})^{\frac{1}{2}}\boldsymbol{\theta}, \sigma^2\mathbf{I}\right) = \frac{1}{(2\pi\sigma^2)^{\frac{N}{2}}}\exp\left[\frac{-1}{2\sigma^2}\left(\mathbf{z}(r,\alpha+\gamma) - (\mathbf{H}^T\mathbf{H})^{\frac{1}{2}}\boldsymbol{\theta}\right)^T \cdot \left(\mathbf{z}(r,\alpha+\gamma) - (\mathbf{H}^T\mathbf{H})^{\frac{1}{2}}\boldsymbol{\theta}\right)\right].$$

The exponent simplifies to $\frac{-1}{2\sigma^2}(r^2 - 2r\boldsymbol{\theta}^T(\mathbf{H}^T\mathbf{H})^{\frac{1}{2}}\begin{bmatrix}\cos(\alpha+\gamma)\\ \sin(\alpha+\gamma)\end{bmatrix} + (\mathbf{H}\boldsymbol{\theta})^T\mathbf{H}\boldsymbol{\theta})$. Under $\mathcal{H}_0$, $p_{\mathbf{z}}(g_\gamma\mathbf{z}(r,\alpha); \mathbf{0}, \sigma^2\mathbf{I}) = \frac{1}{(2\pi\sigma^2)^{\frac{N}{2}}}\exp\left[-\frac{r^2}{2\sigma^2}\right]$. Using these densities in the ratio, where the integration is with respect to Lebesgue measure and since the integrand is continuous over the domain of integration we have the ratio of Riemann integrals, giving

$$\frac{\int_{-\pi}^{\pi} p_{\mathbf{z}}(g_\gamma\mathbf{z}(r,\alpha);\mathcal{H}_1)\,d\gamma}{\int_{-\pi}^{\pi} p_{\mathbf{z}}(g_\gamma\mathbf{z}(r,\alpha);\mathcal{H}_0)\,d\gamma} = \frac{1}{2\pi}\int_{-\pi}^{\pi}\exp\left[\frac{-1}{2\sigma^2}(-2r\boldsymbol{\theta}^T(\mathbf{H}^T\mathbf{H})^{\frac{1}{2}}\begin{bmatrix}\cos(\alpha+\gamma)\\ \sin(\alpha+\gamma)\end{bmatrix} + (\mathbf{H}\boldsymbol{\theta})^T\mathbf{H}\boldsymbol{\theta})\right]d\gamma.$$

Denoting the 1×2 matrix $\boldsymbol{\theta}^T(\mathbf{H}^T\mathbf{H})^{\frac{1}{2}}$ in its polar form by $\begin{bmatrix} r_\beta\cos\beta & r_\beta\sin\beta\end{bmatrix}$ and dropping the non-data dependent term gives

$$\frac{1}{2\pi}\int_{-\pi}^{\pi}\exp\left[\frac{rr_\beta}{\sigma^2}\begin{bmatrix}\cos\beta & \sin\beta\end{bmatrix}\begin{bmatrix}\cos(\alpha+\gamma)\\ \sin(\alpha+\gamma)\end{bmatrix}\right]d\gamma \tag{1}$$

$$= \frac{1}{2\pi}\int_{-\pi}^{\pi}\exp\left[\frac{rr_\beta}{\sigma^2}\cos\gamma'\right]d\gamma' = I_o\left(\frac{rr_\beta}{\sigma^2}\right),$$

where $\gamma' = \beta-\alpha-\gamma$, and $I_o\left(\frac{rr_\beta}{\sigma^2}\right)$ is a modified Bessel function increasing in $\frac{rr_\beta}{\sigma^2}$. Hence, the UMPI test statistic is r (since $\frac{r_\beta}{\sigma^2}$ is positive and not a function of the data), which in terms of the sufficient statistic is equivalent to $\mathbf{z}^T\mathbf{z}$. In terms of the original data samples, we have $\mathbf{z}^T\mathbf{z} = \mathbf{x}^T\mathbf{H}(\mathbf{H}^T\mathbf{H})^{-1}\mathbf{H}^T\mathbf{x}$. Letting $\mathbf{P}_H\mathbf{x} = \mathbf{H}(\mathbf{H}^T\mathbf{H})^{-1}\mathbf{H}^T\mathbf{x}$, which is the projection of $\mathbf{x}$ onto the signal subspace, the statistic can be written as $\mathbf{x}^T\mathbf{P}_H\mathbf{x} = (\mathbf{P}_H\mathbf{x})^T\mathbf{P}_H\mathbf{x}$. The test statistic can be interpreted as the energy of the components of the signal that lie in the signal subspace. This statistic is the GLRT for the linear model. Hence the GLRT is UMPI. Scharf and Friedlander [4] show that the GLRT is UMPI for classes of problems which include this example.

3.1 Relationship between tests

The integral form of Wijsman's theorem gives insight on the relationship between the GLRT and UMPI tests. The form of equation 1, $\int_{-\pi}^{\pi}\exp\left[\frac{rr_\beta}{\sigma^2}\cos\gamma\right]d\gamma$, can be viewed as an 'average over the group.' Let $T(\gamma) = \frac{rr_\beta}{\sigma^2}\cos\gamma$ and let $\hat{\gamma}$ be the value of $\gamma \in [-\pi,\pi]$ that maximizes $T(\gamma)$. Then,

$$\begin{aligned}
&\int_{-\pi}^{\pi}\exp[T(\gamma)]\,d\gamma \\
&= \exp[T(\hat\gamma)]\int_{-\pi}^{\pi}\exp[T(\gamma)-T(\hat\gamma)]\,d\gamma \\
&= \exp\left[\frac{rr_\beta}{\sigma^2}\right]\int_{-\pi}^{\pi}\exp\left[\frac{rr_\beta}{\sigma^2}(\cos\gamma-1)\right]d\gamma.
\end{aligned}$$

The resulting integral is an increasing function of $T(\hat\gamma)$ which is the GLRT for this problem. This integral form suggests the additional interpretation that the GLRT is UMPI since the maximum on the group (GLRT) and the average over the group (UMPI) are 1-1.

Use of this analytic form to investigate the relationship between tests can also be applied to problems for which the integral of the previous equation is not a 1-1 function of the maximum on the group factor. The integral factor may provide insight into the difference between the UMPI and maximum on the group statistics, as well as the conditions under which they become equivalent tests.

Appendix

I Groups

For the lemma described in Section 2.1, the group action must be proper and it is assumed that the groups are locally compact. Consider that for the case of R^N, the action is proper if for any closed and bounded subsets, say $\mathbf{a}$ and $\mathbf{b}$, the set of all group actions such that $\{g : g\mathbf{a} \cap \mathbf{b} \neq \phi\}$ is closed and bounded. In essence, the collection of group elements that transform some point is bounded. As another example consider the scale group G_c (defined below). For any two closed and bounded subsets of R^N, there exists a set $\{c : c\mathbf{a} \cap \mathbf{b} \neq \phi\}$ which is closed and bounded.

It is assumed that the groups are locally compact. Recall that a space in R^N is compact if and only if it is closed and bounded (Heine-Borel theorem). A space is locally compact if and only if every point has at least one compact neighborhood. As an example, on the real line (which is not compact) each point p is interior to a closed interval $[p-\delta, p+\delta]$, which makes it locally compact. A compact space is automatically locally compact and every discrete space is locally compact [13].

A brief description of groups relevant to signal detection applications that meet the conditions of Wijsman's theorem follows.

1. $Al_n = \{g_{A,b} : g_{A,b}\mathbf{x} = \mathbf{A}\mathbf{x} + \mathbf{b}\}$, where $\mathbf{A}$ are N x N nonsingular matrices and $\mathbf{b} \in R^N$. This is known as the affine group. This group is not compact. It is locally compact [9], and the group action is proper [7].

2. $Gl_n = \{g_A : g_A\mathbf{x} = \mathbf{A}\mathbf{x}\}$, where $\mathbf{A}$ are N x N nonsingular matrices. This is the general linear group. This is a subgroup of the affine group. The group is not compact, is locally compact [9] and acts properly [7].

3. $R^n = \{g_b : g_b\mathbf{x} = \mathbf{x} + \mathbf{b},\ \mathbf{b} \in R^N\}$ is a subgroup of Al_n. This group is not compact, is locally compact [9], and acts properly [7]. A subgroup of this is the translation group, where $\mathbf{b} = b\mathbf{1}$, which is locally compact and acts properly.

4. $G_c = \{g_c : g_c\mathbf{x} = c\mathbf{x},\ c \in R^1\backslash 0\}$ is the scale group. It is a subgroup of Gl_n. It is not compact, is locally compact. Proper action is inherited from Gl_n.

5. $G_s = \{g_s : g_s\mathbf{x} = c\mathbf{x},\ c \in \{-1, 1\}\}$ is the sign group. It is a subgroup of G_c. It is discrete and hence is compact, locally compact, and acts properly since discrete groups are compact and act properly [7].

6. $G_o = \{g_o : g_o\mathbf{x} = \mathcal{O}\mathbf{x},\ \mathcal{O}^T\mathcal{O} = \mathbf{I}\}$ is the orthogonal group. It is a subgroup of Gl_n. It is compact and hence is locally compact and acts properly.

7. $P_n = \{g_k : g_k\mathbf{x} = \mathbf{P}^k\mathbf{z}\}$, where $\mathbf{P}^k$ are the N x N permutation matrices. These contain rows and columns with exactly one element equal to 1 and the others are 0. This group is discrete and hence is compact, locally compact and acts properly.

II Invariance of Examples

In example 1, the problem is invariant under G_s since $p_{\mathbf{x}}\left(g_s^{-1}\mathbf{x}; A\mathbf{s}\right)\left|J_{g_s^{-1}}(\mathbf{x})\right|$ is

$$\begin{aligned}
&\frac{1}{(2\pi\sigma^2)^{\frac{N}{2}}}\exp\left[\frac{-1}{2\sigma^2}\sum_{n=0}^{N-1}\left(g_s^{-1}x[n] - As[n]\right)^2\right] \\
=\ &\frac{1}{(2\pi\sigma^2)^{\frac{N}{2}}}\exp\left[\frac{-1}{2\sigma^2}\sum_{n=0}^{N-1}\left(x[n] - \bar{g}_s As[n]\right)^2\right] \\
=\ &p_{\mathbf{x}}\left(\mathbf{x}; \bar{g}_s A\mathbf{s}\right),
\end{aligned}$$

since $g_s^{-1} = \bar{g}_s$. The Jacobian $\left|J_{g_s^{-1}}(\mathbf{x})\right|$ of the transformation is 1. Hence, transformation of $\mathbf{x}$ by these group elements results in a random variable belonging to the same family of densities, with a different value of the parameter that belongs to the same hypothesis.

In example 2, the problem is invariant under G_o since under $\mathcal{H}_1$

$$\begin{aligned}
&p_{\mathbf{z}}\left(g_o^{-1}\mathbf{z}; \left(\mathbf{H}^T\mathbf{H}\right)^{\frac{1}{2}}\boldsymbol{\theta}, \sigma^2\mathbf{I}\right)\left|J_{g_o^{-1}}(\mathbf{z})\right| \\
&= \frac{1}{(2\pi\sigma^2)^{\frac{N}{2}}}\exp\left[\begin{array}{c}\frac{-1}{2\sigma^2}\left(\mathcal{O}^T\mathbf{z} - \left(\mathbf{H}^T\mathbf{H}\right)^{\frac{1}{2}}\boldsymbol{\theta}\right)^T \\ \cdot\left(\mathcal{O}^T\mathbf{z} - \left(\mathbf{H}^T\mathbf{H}\right)^{\frac{1}{2}}\boldsymbol{\theta}\right)\end{array}\right] \\
&= \frac{1}{(2\pi\sigma^2)^{\frac{N}{2}}}\exp\left[\begin{array}{c}\frac{-1}{2\sigma^2}\left(\mathbf{z} - \mathcal{O}\left(\mathbf{H}^T\mathbf{H}\right)^{\frac{1}{2}}\boldsymbol{\theta}\right)^T \\ \cdot\left(\mathbf{z} - \mathcal{O}\left(\mathbf{H}^T\mathbf{H}\right)^{\frac{1}{2}}\boldsymbol{\theta}\right)\end{array}\right] \\
&= p_{\mathbf{z}}\left(\mathbf{z}; \bar{g}_o\left(\mathbf{H}^T\mathbf{H}\right)^{\frac{1}{2}}\boldsymbol{\theta}, \sigma^2\mathbf{I}\right).
\end{aligned}$$

After transformation, the PDF belongs to the same family of PDF's with a value of the parameter belonging to the same hypothesis. Similarly, under $\mathcal{H}_0$, $p_{\mathbf{z}}\left(g_o^{-1}\mathbf{z}; \mathbf{0}, \sigma^2\mathbf{I}\right)\left|J_{g_o^{-1}}(\mathbf{z})\right| = p_{\mathbf{z}}\left(\mathbf{z}; \mathbf{0}, \sigma^2\mathbf{I}\right)$.

III Multiplier and left invariant measure for the scale group

This appendix provides an additional example of both the invariant measure and multiplier. The scale group is an interesting example since the multiplier is not 1 and the invariant measure is not simply a counting measure or the measure of an interval.

Consider the transformation $\mathbf{y} = g_c\mathbf{x} = c\mathbf{x}$, $c \neq 0$. Assuming Gaussian densities with unknown mean $\mu\mathbf{1}$ and variance $\sigma^2\mathbf{I}$, the multiplier can be obtained from the following PDF relationship,

$$\begin{aligned} & p_{\mathbf{x}}\left(g_c^{-1}\mathbf{x};\mu\mathbf{1},\sigma^2\mathbf{I}\right)\left|J_{g_c^{-1}}\left(\mathbf{x}\right)\right| \\ = & \frac{1}{\left(2\pi\sigma^2\right)^{\frac{N}{2}}}\exp\left[\frac{-1}{2\sigma^2}\left(\frac{\mathbf{x}}{c}-\mu\mathbf{1}\right)^T\left(\frac{\mathbf{x}}{c}-\mu\mathbf{1}\right)\right]|c|^{-N} \\ = & \frac{1}{\left(2\pi c^2\sigma^2\right)^{\frac{N}{2}}}\exp\left[\frac{-1}{2c^2\sigma^2}\left(\mathbf{x}-c\mu\mathbf{1}\right)^T\left(\mathbf{x}-c\mu\mathbf{1}\right)\right]. \end{aligned}$$

Hence (see Section II, Jacobian and multiplier discussion) the multiplier is $\chi_0\left(g_c\right) = \left|J_{g_c^{-1}}\left(\mathbf{x}\right)\right|^{-1} = |c|^N$.

A left invariant measure for scale group is described next. Consider that for this group, the composition of two group elements is $g_c g_{c_1}\mathbf{x} = g_{cc_1}\mathbf{x} = cc_1\mathbf{x}$, and this composite action occurs in R^1. An invariant measure in R^1 with $A = [a_1, a_2] \subset R^1$ should satisfy $\nu_l\left(g_c A\right) = \nu_l\left(A\right)$. Note that the Lebesgue measure $\int_A dx = a_2 - a_1 \neq \int_{cA} dx$ is not invariant. We need an invariant measure, $\nu_l\left(cA\right) = \nu_l\left(A\right) \forall c \neq 0$. For ν_l absolutely continuous with respect to a Lebesgue measure, then by the Radon-Nikodym theorem, $\nu_l\left(A\right) = \int_A g\left(x\right)dx$, where $g\left(x\right)$ is a unique nonnegative measurable function. Hence, for $\nu_l\left(A\right)$ to be invariant, it must equal $\nu_l\left(cA\right) = \int_{cA} g\left(x\right)dx$. Let $u = \frac{x}{c}$. Then $\nu_l\left(cA\right) = \int_A g\left(cu\right)|c|\,du = \nu_l\left(A\right)$. This implies that we need $g\left(cu\right)|c| = g\left(u\right)$. A solution is $g\left(u\right) = \frac{1}{|u|}$, where we define $g\left(0\right) = 0$. Hence, $\nu_l\left(A\right) = \int_A \frac{1}{|u|}du$ which in the notation of Wijsman's theorem is $\frac{d\nu_l(c)}{dc} = \frac{1}{|c|}$, or $d\nu_l\left(c\right) = \frac{1}{|c|}dc$.

For the problem of detection of a known signal of unknown level in white Gaussian noise of unknown variance, the substitutions into the theorem gives $\frac{\int p_{\mathbf{x}}(c\mathbf{x};\mathcal{H}_1)|c|^{N-1}dc}{\int p_{\mathbf{x}}(c\mathbf{x};\mathcal{H}_0)|c|^{N-1}dc}$, which is easily simplified into the test statistic $\left|\frac{\mathbf{s}^T\mathbf{x}}{\sqrt{\mathbf{x}^T\mathbf{x}}}\right|$.

References

[1] Steven M. Kay, *Fundamentals of Statistical Signal Processing: Detection Theory*, vol. II, Prentice-Hall PTR, Upper Saddle River, New Jersey, 1998.

[2] E.L. Lehmann, *Testing Statistical Hypotheses*, Springer-Verlag, New York, 2nd edition, 1986.

[3] Louis L. Scharf, *Statistical Signal Processing: Detection, Estimation, and Time Series Analysis*, Addison-Wesley, 1991.

[4] Louis L. Scharf and Benjamin Friedlander, "Matched subspace detectors," *IEEE Transactions on Signal Processing*, vol. 42, no. 8, pp. 2146–2157, August 1994.

[5] C. Stein, "Some problems in multivariate analysis, part 1.," *Technical Report No. 6, Department of Statistics, Stanford University*, 1956.

[6] Robert A. Wijsman, "Cross-sections of orbits and their application to densities of maximal invariants," *Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability*, vol. 1, pp. 389–400, 1967.

[7] Robert A. Wijsman, "Proper action in steps, with application to density ratios of maximal invariants," *The Annals of Statistics*, vol. 13, pp. 395–402, 1985.

[8] S. Andersson, "Distributions of maximal invariants using quotient measures," *The Annals of Statistics*, vol. 10, pp. 955-961, 1982.

[9] Morris L. Eaton, *Group Invariance Applications in Statistics*, Institute of Mathematical Statistics and the American Statistical Association, 1989.

[10] Vincent A. Uthoff, "The most powerful scale and location invariant test of the normal versus the double exponential," *The Annals of Statistics*, vol. 1, pp. 170–174, 1973.

[11] Takeaki Kariya and Bimal K. Sinha, *Robustness of Statistical Tests*, Academic Press, San Diego, CA, 1989.

[12] Richard E. Schwartz, "Minimax CFAR detection in additive Gaussian noise of unknown covariance," *IEEE Transactions on Information Theory*, pp. 722–725, November 1969.

[13] John L. Kelley, *General Topology*, Springer-Verlag, New York, 1975.