Constrained Inference for Double Cone Alternatives

Xuechan Li* Janice McCarthy* Zhiguo Li* Andrew Allen*
Kouros Owzar*

Abstract

In medical studies, prior knowledge can often be used to constrain inference to clinically or biologically relevant alternative hypotheses, leading to substantial power gains relative to an unconstrained approach. For example, in cancer pharmacogenomics studies, researchers may be interested in markers for which the gene effect is present only when exposed to drug. Often the space of interesting alternatives can be described by the boundary or closure of a double cone. While closed convex cone alternatives have been well-studied, previous studies of closed double cone alternatives have been limited to empirical investigations of the type I error. Here we present a detailed treatment of inference for double cone alternatives. We derive explicit mathematical formulas for calculating type I and type II error rates and illustrate how these rates relate to geometric features of the acceptance region. We provide numerical algorithms for approximating the error rates and evaluate their performance through simulations studies.

Key Words: constrained inference, double cone, geometry, likelihood ratio test

1. Introduction

In statistical study, researchers often have clear expectations about the direction of the parameters in their statistical model. For example, in a cancer pharmacogenomics study, researchers may only interested in markers for which the gene effect is absent when unexposed to drug or when exposed to drug. Thus the parameter space need not to be the whole Euclidean space. The alternative region by taking the additional knowledge into account might introduce larger hypothesis testing power.

The restricted alternative region can be closed convex cone or closed double cone. Where **closed convex cone** $\mathcal{C}^s \in \mathbb{R}^p$ is defined as "A convex set that consists of infinite straight lines starting from the origin" [Silvapulle and Sen [2011]]. Thus if $\mathbf{x}, \mathbf{y} \in \mathcal{C}^s$, then $\lambda_1\mathbf{x}+\lambda_2\mathbf{y} \in \mathcal{C}^s$, for all $\lambda_1 \geq 0$, $\lambda_2 \geq 0$. The **closed double cone** $\mathcal{C} \in \mathbb{R}_p$ in this paper is the set such that if $\mathbf{x}, \mathbf{y} \in \mathcal{C}$, then $\lambda_1\mathbf{x}+\lambda_2\mathbf{y} \in \mathcal{C}^s$, for all $\lambda_1, \lambda_2 \in \mathbb{R}$. **Figure 1** illustrates the closed convex cone and closed double cone in two dimensional parameter space.

There is an extensive body of literature on the single closed, convex cone statistical inference. Tsai [1992] shows that for a multivariate normal vector X with known covariance matrix σ, the restricted likelihood ratio test (LRT) for $H_0 : \mu = 0$ versus the orthant alternative $H_A : \mu \leq 0$ is uniformly more powerful than the test against $H_A : \mu \in \mathbb{R}^p$. Praestgaad [2012] studies the power superiority in a more general situation. He proved that the restricted LRT for $H_0 : \mu \in \mathcal{L}$ versus $H_A : \mu \in \mathcal{C}^s/\mathcal{L}$ is uniformly more powerful

*Duke University, 2424 Erwin Road, Durham, NC 27705

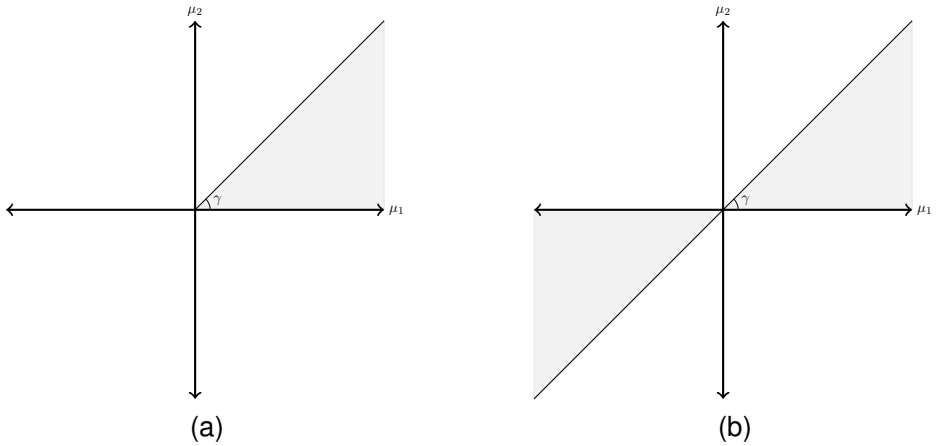


Figure 1: The figure showing the closed convex cone and double cone. Where the shaded region in (a) illustrate the closed convex cone and the shaded region in (b) illustrate the closed double cone. γ is the vertex angle for both closed convex cone and closed double cone. In (a), we test $H_0 : \mu_1 = \mu_2 = 0$ versus $H_A : \mu_2 \geq 0, \mu_2 - \mu_1 \tan\gamma \geq 0$; in (b), we test $H_0 : \mu_1 = \mu_2 = 0$ versus $H_A : \mu_2(\mu_2 - \mu_1 \tan\gamma) \leq 0$

than the unconstrained test. At here, $\mathcal{L}$ is a linear subspace of the closed convex cone $\mathcal{C}^s$.

As for the power function. Hu and Wright [1994] shown that for a multivariate normal model with known covariance matrix, the restricted LRT for $H_0 : \mu \in \mathcal{L}$ versus $H_A : \mu \in \mathcal{C}^s/\mathcal{L}$ will have non-decreasing power on each line segment in $\mathcal{C}^s$ starting at a point in $\mathcal{L}$. Iwasa [1998] proves the power functions' directional monotonicity properties for closed convex cone restricted hypothesis of normal means. A partial ordering "$\leq_{\mathcal{C}^s}$" has been defined in the paper, and the author proved that the power function is increasing in $\leq_{\mathcal{C}^s}$.

The closed convex cone restricted LRT's asymptotic distribution is a mixture of chi-square distribution [Silvapulle and Sen, 2011]. To study the effect of the weight on the critical points and the power function, Nüesch [1966] provided an numerical study for the weight effect on the critical points and the power function in multivariate normal model. Gourieroux et al. [1982] also provides the numerical study for the weight effect on the critical points in linear model.

There are much less study for the closed double cone restricted LRT. Although it also has a bright prospect to be applied in medical research. Sen and Meyer [2016] developed the restricted LRT under multi-dimensional double cone alternatives. They computed the critical points for fixed level α via simulation and shown that the power of the test will converge to 1 when sample size increases. McCarthy et al.[2014] developed the closed double cone restricted LRT and score test to test for the genetic effects related to HIV-1 acquisition. Since when considering the HIV-1, HRSN (high-risk seronegatives), and population control samples together, researcher are clear that risk alleles will neither enriched nor depleted in both HIV-1 and HRSN group. The parameter space for a gene's effects on two diseases will be a double cone. They find that the constrained testing approach will have about 10-28% power gain compared to the typical unconstrained test.

In the McCarthy et al's paper, they shown that the distribution for the closed double cone restricted LRT is still a mixture of distribution. Thus the distribution is a weighted sum of different distribution. Nüesch [1966] also provide an numerical treatment for the weight function's effect on the critical points under the closed double cone restricted alternatives. To the best of our knowledge, there is no following study for weight effect on the critical points and power under this case.

In two dimensional parameter space, there is an surjective mapping between cone's vertex angle and the weight function. Our study will focus on the bivariate normal distribution with identity covariance matrix.

In this paper, we consider inference on two types of restricted alternative regions: a closed double cone and its boundary. First, we provide a detailed outline of the two hypotheses, including a geometric representation of two corresponding regions. For a fixed type I error probability, we establish the relationship between the vertex angle of the double cone with the critical point of the likelihood ratio test and its power. These relationships are formulated both analytically and geometrically. We compare the power of the restricted tests to that of an unrestricted test. While the discussions are presented in the context of standard bivariate normal distribution, we will outline the extension to an asymptotically bivariate normal distribution with an arbitrary positive definite covariance matrix. Finally, we discuss future extensions.

2. Method

2.1 Model and Hypothesis

The observation is drawn from $N(\mu, \mathbb{I})$, a bivariate normal distribution with mean $\mu = (\mu_1, \mu_2)^T$ and identity variance. We are interested in testing the null hypothesis $H_0 : \mu = (0,0)^T$ versus the alternative $H_1 : \mu \in \mathcal{C}$, where $\mathcal{C} \subset \mathbb{R}^2$. We are specifically interested in the case where $\mathcal{C}$ is a closed double cone, to be denoted by $\mathcal{C}_1$, and the boundary of closed double cone, to be denoted by $\mathcal{C}_2$. We can define these two sets as

$$\mathcal{C}_1 = \{\mu \in \mathbb{R}^2 : (\mu_2 - \mu_1 \tan\gamma_1)(\mu_2 - \mu_1 \tan\gamma_2) \leq 0\},$$

and

$$\mathcal{C}_2 = \{\mu \in \mathbb{R}^2 : (\mu_2 - \mu_1 \tan\gamma_1)(\mu_2 - \mu_1 \tan\gamma_2) = 0\}$$

respectively, where $0 \leq \gamma_1 \leq \gamma_2 \leq \pi$. In the two dimensional space, the boundaries for both $\mathcal{C}_1$ and $\mathcal{C}_2$ are two straight lines. Since the normal distribution with identity variance is rotationally invariant, we can simplify the problem by rotating the distribution and alternative region, so that one of the boundaries of the alternative region will always be the x-axis (**Figure 2**). The region for the $\mathcal{C}_1$ and $\mathcal{C}_2$ can be presented as

$$\mathcal{C}_1 = \{(\mu_1, \mu_2)^T \in \mathbb{R}^2 : \mu_2(\mu_2 - \mu_1 \tan\gamma) \leq 0\},$$

where $\gamma \in [0, \pi]$, and

$$\mathcal{C}_2 = \{\mu \in \mathbb{R}^2 : \mu_2(\mu_2 - \mu_1 \tan\gamma) = 0\},$$

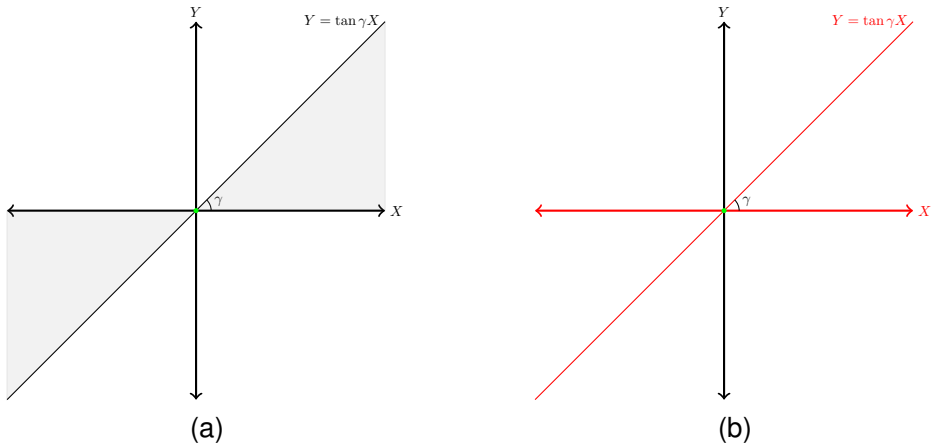


Figure 2: Illustration of the parameter spaces of hypotheses 1 and 2. The shaded region in panel (a) corresponds to the double cone shaped parameter space $\mathcal{C}_1$, for hypothesis 1. The red lines in panel (b) correspond to the parameter space $\mathcal{C}_2$ for hypothesis 2. In each case, the null space is denoted by a green dot.

where $\gamma \in [0, \pi/2]$. Here $\gamma = |\gamma_2 - \gamma_1|$ is the vertex angle of the double cone.

2.2 Test Statistics and Distribution

We define the likelihood ratio statistic for testing $H_0 : \mu \in \mathcal{C}_0 = (0,0)^T$ versus $H_1 : \mu \in \mathcal{C}$ as $T = 2\big[\max_{\mu\in\mathcal{C}} \ell(\mu) - \max_{\mu\in\mathcal{C}_0} \ell(\mu)\big]$, where $\ell(\mu)$ denotes the log-likelihood function. For the maximum likelihood estimator (MLE) restricted to $\mathcal{C}$, $\tilde{\mu} = (\tilde{\mu}_1, \tilde{\mu}_2)^T = \arg\max_{\mu\in\mathcal{C}} \ell(\mu)$, we have that

$$\begin{aligned} T &= \|(X,Y)^T\|^2 - \|(X-\tilde{\mu}_1, Y-\tilde{\mu}_2)^T\|^2 \\ &= \|(\tilde{\mu}_1,\tilde{\mu}_2)^T\|^2 + 2(X-\tilde{\mu}_1, Y-\tilde{\mu}_2)(\tilde{\mu}_1,\tilde{\mu}_2)^T. \end{aligned}$$

If $((X,Y)^T \notin \mathcal{C})$, then $(\tilde{\mu}_1,\tilde{\mu}_2)^T$ is on the boundary of the restricted region $\mathcal{C}$. Otherwise, if $(X,Y)^T \in \mathcal{C}$, then $(\tilde{\mu}_1,\tilde{\mu}_2)^T = (X,Y)^T$. In either case, as $(\tilde{\mu}_1,\tilde{\mu}_2)^T$ is a projection of $(X,Y)^T$ onto the restricted region $\mathcal{C}$, $(X-\tilde{\mu}_1, Y-\tilde{\mu}_2)(\tilde{\mu}_1,\tilde{\mu}_2)^T = 0$. Consequently, $T = \|(\tilde{\mu}_1,\tilde{\mu}_2)^T\|^2$, which implies that the restricted likelihood ratio statistics is the square of the length of the projection of $(X,Y)^T$ onto $\mathcal{C}$.

For the unconstrained case, where $\mathcal{C} = \mathbb{R}^2$, the test statistic is trivially observed to be $T_0 = X^2 + Y^2$. Under the null, it follows a chi-square distribution with two degrees of freedom to be denoted by χ^2_2. Recalling that the boundary of a double cone in $\mathbb{R}^2$ consists of two straight lines, we let U and V denote the projections of $(X,Y)^T$ onto these two lines respectively. Then the square of the length of the projections of $(X,Y)^T$ onto $\mathcal{C}$ is $\max\{\|U\|^2, \|V\|^2\}$.

For hypothesis 1, the test statistic is

$$T_1 = (X^2+Y^2)\mathbb{1}\{(X,Y)^T \in \mathcal{C}_1\} + \max\{\|U\|^2, \|V\|^2\}(1 - \mathbb{1}\{(X,Y)^T \in \mathcal{C}_1\}).$$

Its null sampling distribution is given by

$$\Pr(T_1 \leq c) = (\frac{\gamma}{\pi}) \Pr(W \leq c) + (1 - \frac{\gamma}{\pi}) \Pr(\|U\|^2 \leq c, \|V\|^2 \leq c | (X, Y) \notin \mathcal{C}_1),$$

where W is distributed according to χ^2_2. For hypothesis 2, the test statistic is $T_2 = \max\{\|U\|^2, \|V\|^2\}$. Its sampling distribution is

$$\Pr(T_2 \leq c) = \Pr(\|U\|^2 \leq c, \|V\|^2 \leq c | (X, Y) \notin \mathcal{C}_2) = \Pr(\|U\|^2 \leq c, \|V\|^2 \leq c).$$

As $(U, V)^T$ is a linear transformation of $(X, Y)^T$ it follows a bivariate normal distribution.

3. Angle Effect for Critical Point and Power

In this section, we conduct numerical studies of the angle effect on both the critical point and power for hypotheses 1 and 2. The integrals are approximated numerically using the R package cubature [Johnson and Narasimhan [2013]]. The type I error probability and power are approximated using $100,000$ Monte Carlo replicates.

3.1 Critical Point

We defined the critical point r as the maximum acceptable length of the projection of observed point $(x, y)^T$ onto the restricted region. The acceptance regions for hypothesis 1 and 2 are denoted by $G_1 = G_1(r, \gamma)$ and $G_2 = G_2(r, \gamma)$, respectively. As shown in **Figure 3**, G_1 is a compass needle shaped region while G_2 is a diamond shaped region. For a fixed angle γ, the size for the acceptance region

$$\mu(G_k) = \int_{G_k(\gamma, r)} \frac{1}{2\pi} \exp\{-\frac{x^2 + y^2}{2}\} dxdy,$$

, where $k \in \{1, 2\}$, is an increasing function in r.

Here, we fix the type I error $\alpha = 0.05$ and angle γ, and then numerically solve the equation $\Pr(G_k(r, \alpha)) = 1 - \alpha$ for the critical point r using bisection. The mathematical details to show that for a fixed α, r is increases as the angle γ increased are provided in **Appendix A**. The results are shown in **Table 1** for hypotheses 1 and 2. As noted previously, for hypothesis 1, the angle ranges from 0 to π while for hypothesis 2, it is bounded above by $\frac{\pi}{2}$. As expected, for both hypotheses, the critical point r increases as the angle γ increases.

3.2 Power

For hypotheses 1 and 2, the power can be presented as

$$\text{power} = 1 - \int_{G_k} f(x - \mu_1, y - \mu_2) dxdy = 1 - \int_{G_k} \frac{1}{2\pi} \exp\{-\frac{(x - \mu_1)^2 + (y - \mu_2)^2}{2}\} dxdy,$$

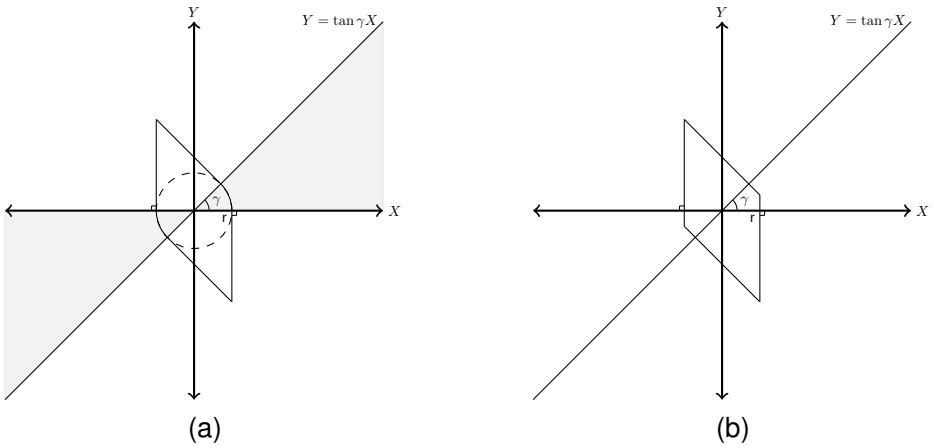


Figure 3: This figure illustrates the acceptance region for hypotheses 1 and 2. The compass needle shaped region in panel (a) corresponds to the acceptance region G_1 while the diamond shaped region in panel (b) corresponds to the acceptance region G_2.

Table 1: Critical points r and empirical type I error for hypothesis 1 and 2. For a given angle, the critical point r is approximated numerically using bisection at the nominal $\alpha = 0.05$ level. Next, for the given angle and approximated critical point r, the type I error is confirmed empirically through simulation ($\hat{\alpha}$).

Hypothesis		Angle (γ)						
		0	$\pi/6$	$\pi/3$	$\pi/2$	$2\pi/3$	$5\pi/6$	π
Hypothesis 1	r	1.960	2.134	2.258	2.352	2.415	2.443	2.448
	$\hat{\alpha}$	0.0498	0.0496	0.0498	0.0497	0.0504	0.0502	0.0498
Hypothesis 2	r	1.960	2.126	2.212	2.236	-	-	-
	$\hat{\alpha}$	0.0498	0.0503	0.0497	0.0499	-	-	-

where $k \in \{1, 2\}$. The numerical and empirical study on power is conducted on different distances: $d \in \{1, 2.5, 4\}$. For hypothesis 1, we assume the true parameter $(\mu_1, \mu_2) = (d\cos(\frac{\gamma}{2}), d\sin(\frac{\gamma}{2}))$, where $\gamma \in \{0, \pi/6, \pi/3, \pi/2, 2\pi/3, 5\pi/6, \pi\}$. Note that the restricted test with $\gamma = \pi$ is actually the unrestricted test in hypothesis 1. As for hypothesis 2, the study is on different angles: $\gamma \in \{0, \pi/6, \pi/3, \pi/2\}$. We also assume the true parameter $(\mu_1, \mu_2) = (d, 0)$.

Table 2 and **Table 3** show the numerical study results for hypothesis 1 and 2, respectively. We observe that the power is decreasing when γ increasing under both of the hypothesis 1 and 2 when the distance is fixed. Comparing to the unrestricted test, the restricted test will always have larger power than the unconstrained test in hypothesis 1. While for the hypothesis 2, restricted test is more powerful in most of the scenario. The only exception appeals when distance $d = 1$ and angle $\gamma = \frac{\pi}{2}$, where the unrestricted test is slightly more powerful than the restricted test.

Table 2: Power for hypothesis 1. For a fixed distance $d \in \{1, 2.5, 4\}$, the power is approximated numerically as a function of the angle at the nominal $\alpha = 0.05$ level. The approximated critical points from Table 1 are used.

Distance (d)	Angle (γ)							
	0	$\pi/6$	$\pi/3$	$\pi/2$	$2\pi/3$	$5\pi/6$	π	Unconstrained
1	0.170	0.167	0.158	0.147	0.138	0.133	0.133	0.133
2.5	0.705	0.695	0.669	0.639	0.616	0.605	0.603	0.603
4	0.979	0.977	0.971	0.965	0.960	0.957	0.957	0.957

Table 3: Power for hypothesis 2. For a fixed distance $d \in \{1, 2.5, 4\}$, the power is approximated numerically as a function of the angle at the nominal $\alpha = 0.05$ level. The approximated critical points from Table 1 are used.

Distance (d)	Angle (γ)				
	0	$\pi/6$	$\pi/3$	$\pi/2$	Unconstrained
1	0.170	0.160	0.141	0.131	0.133
2.5	0.705	0.678	0.634	0.614	0.603
4	0.979	0.974	0.966	0.962	0.957

4. Discussion

In this report, we have studied the angle effect on the type I and type II error rates for two types of constrained hypotheses. For both cases, we show that as the angle γ increases, the critical point r will increase (to the maximum point) and illustrate these properties through numerical examples. For these examples, we observe that the power decreases as the angle increases and that the restricted test for hypothesis 1 has larger power than the unrestricted test. For hypothesis 2, under most scenarios, the restricted test has larger power than the unrestricted test.

Our study is based on a bivariate normal distribution with identity variance. We can easily, as shown in **Appendix B**, extend our study to a asymptotically bivariate normal distribution with arbitrary positive definite covariance matrix. More effort is needed to extend the study to higher dimensions. More specifically, a new geometric feature in stead of the angle needs to be found. Furthermore, in a two-dimensional parameter space, the closed double cone will always be the circular cone, but in higher dimensions, the closed double cone can be circular, polyhedral or other types.

REFERENCES

Gourieroux, C., Holly, A., and Monfort, A. (1982), "Likelihood ratio test, Wald test, and Kuhn-Tucker test in linear models with inequality constraints on the regression parameters," *Econometrica: Journal of the Econometric Society*, 63-80.

Hu, X., and Wright, F. T. (1994), "Monotonicity properties of the power functions of likelihood ratio tests for normal mean hypotheses constrained by a linear space and a cone," *The Annals of Statistics*, 1547-1554.

Iwasa, M. (1998), "Directional monotonicity properties of the power functions of likelihood ratio tests for cone-restricted hypotheses of normal means," *Journal of Statistical Planning and Inference*, 66(2):223-233.

Johnson, S. G., and Narasimhan, B. (2013), "cubature: Adaptive multivariate integration over hypercubes," URL https://CRAN.R-project.org/package=cubature. R package version 1.1-2.

McCarthy, J. M., Shea, P. R., Goldstein, D. B., and Allen, A. S. (2014), "Testing for risk and protective trends in genetic analyses of HIV acquisition," *Biostatistics*, kxu044.

Nüesch, P. E. (1966), "On the problem of testing location in multivariate populations for restricted alternatives," *The Annals of Mathematical Statistics*, 37(1):113-119.

Praestgaard, J. (2012), "A note on the power superiority of the restricted likelihood ratio test," *Journal of Multivariate Analysis*, 104(1):1-15.

Sen, B., and Meyer, M. (2016), "Testing against a linear regression model using ideas from shape-restricted estimation," *Journal of the Royal Statistical Society*, Ser. B (Statistical Methodology).

Silvapulle, M. J., and Sen, P. K. (2011), *Constrained Statistical Inference: Order, inequality, and shape constraints* (volume 912.), John Wiley & Sons.

Tsai, M. (1992), "On the power superiority of likelihood ratio tests for restricted alternatives," *Journal of Multivariate Analysis*, 42(1):102-109.

A. Relationship between Vertex Angle and Critical Points

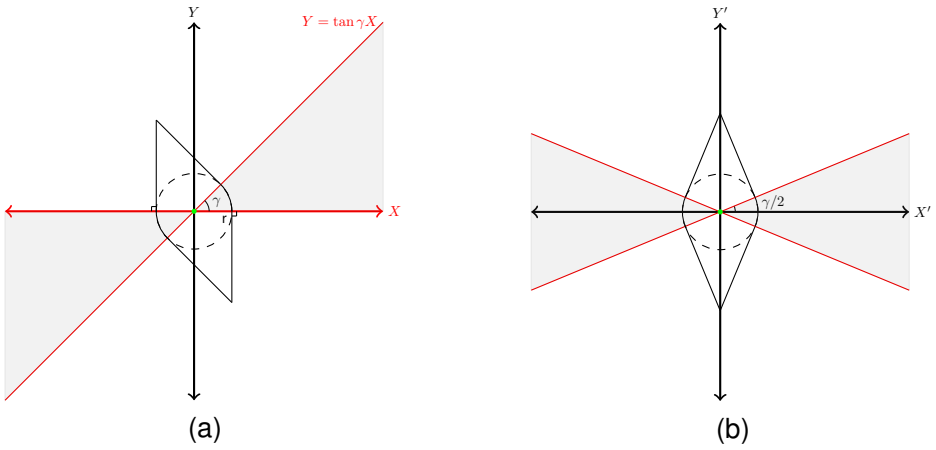
In this section, we provide a mathematical study of the angle impact on critical point r is provided. For both hypotheses 1 and 2, we prove that, for a fixed significance level α, the critical point r is an increasing function of the angle γ, where $\gamma \in [0, \pi]$ for hypothesis 1 and $\gamma \in [0, \frac{\pi}{2}]$ for hypothesis 2.

As previously defined G_1 and G_2 denote the acceptance regions for hypotheses 1 and 2. The size of the acceptance region, under H_0, is then obtained as

$$1 - \alpha = \mu(G_k) = \int_{G_k} \frac{1}{2\pi} \exp\{-\frac{x^2 + y^2}{2}\} dx dy,$$

where the k=1,2.

A.1 Hypothesis 1



(a) (b)

Figure 4: The compass needle shaped region G_1 in (a) is the acceptance region for hypothesis 1 under significance level $1 - \mu(G_1)$. The diamond shaped region in panel (b) is obtained by rotating the region in panel (a) through the angle $\frac{\gamma}{2} \in [0, \frac{\pi}{2}]$ in a clockwise direction.

Based on the rotational symmetry of the standard bivariate normal distribution, when γ and r are fixed, the diamond regions in panels (a) and (b) in (**Figure 4**) have same Gaussian measure. Thus it is sufficient to prove that the critical point r is increasing when the angle $\beta = \frac{\gamma}{2}$ increases under the fixed Gaussian measure of the region in panel (b).

Here, we use polar coordinates to simplify the calculation. For a given point (x, y) in panel (a), we represent the corresponding point in panel (b) as $(x', y') = (\rho\cos\theta, \rho\sin\theta)$, for some $\rho \in \mathbb{R}$ and $\theta \in [0, 2\pi]$. In terms of these polar coordinates, when the significant

level is fixed at $\alpha \in [0,1]$, we have:

$$\begin{aligned}\frac{\pi}{2}(1-\alpha) =& \int_0^{\beta}\int_0^{r} \exp\{-\frac{\rho^2}{2}\}\rho d\rho d\theta + \int_{\beta}^{\frac{\pi}{2}}\int_0^{\frac{r}{\cos(\theta-\beta)}} \exp\{-\frac{\rho^2}{2}\}\rho d\rho d\theta \\ =& \beta(1-\exp\{-\frac{r^2}{2}\}) + \int_{\beta}^{\frac{\pi}{2}} 1-\exp\{-\frac{r^2}{2\cos^2(\theta-\beta)}\}d\theta.\end{aligned}$$

To find $r' = \frac{dr}{d\beta}$ implicitly, we first differentiate both sides of the equation above to get

$$\begin{aligned}0 =& \frac{d}{d\beta}\left[\beta(1-\exp\{-\frac{r^2}{2}\}) + \int_{\beta}^{\frac{\pi}{2}} 1-\exp\{-\frac{r^2}{2\cos^2(\theta-\beta)}\}d\theta\right] \\ =& \beta r\exp\{-\frac{r^2}{2}\}r' - \int_{\beta}^{\frac{\pi}{2}} \frac{d}{d\beta}\exp\{-\frac{r^2}{2\cos^2(\theta-\beta)}\}d\theta \\ =& \left[\beta r\exp\{-\frac{r^2}{2}\} + \int_{\beta}^{\frac{\pi}{2}} \exp\{-\frac{r^2}{2\cos^2(\theta-\beta)}\}\frac{r}{\cos^2(\theta-\beta)}d\theta\right]r' \\ & - \int_{\beta}^{\frac{\pi}{2}} \exp\{-\frac{r^2}{2\cos^2(\theta-\beta)}\}\frac{r^2\sin(2(\theta-\beta))}{2\cos^4(\theta-\beta)}d\theta.\end{aligned}$$

When $\theta \in (\beta, \frac{\pi}{2})$, we have $\sin(2(\theta-\beta)) > 0$ and so

$$\int_{\beta}^{\frac{\pi}{2}} \exp\{-\frac{r^2}{2\cos^2(\theta-\beta)}\}\frac{r^2\sin(2(\theta-\beta))}{2\cos^4(\theta-\beta)}d\theta > 0.$$

Since it is trivial to find

$$\beta r\exp\{-\frac{r^2}{2}\} + \int_{\beta}^{\frac{\pi}{2}} \exp\{-\frac{r^2}{2\cos^2(\theta-\beta)}\}\frac{r}{\cos^2(\theta-\beta)}d\theta > 0,$$

we get

$$r' = \frac{\int_{\beta}^{\frac{\pi}{2}} \exp\{-\frac{r^2}{2\cos^2(\theta-\beta)}\}\frac{r^2\sin(2(\theta-\beta))}{2\cos^4(\theta-\beta)}d\theta}{\beta r\exp\{-\frac{r^2}{2}\} + \int_{\beta}^{\frac{\pi}{2}} \exp\{-\frac{r^2}{2\cos^2(\theta-\beta)}\}\frac{r}{\cos^2(\theta-\beta)}d\theta} > 0$$

Thus the critical point r is an increasing function in $\beta = \frac{\gamma}{2} \in [0, \frac{\pi}{2}]$ and consequently in $\gamma \in [0, \pi]$.

A.2 Hypothesis 2

Similar to the discussions for Hypothesis 1, it is sufficient for us to show that r is an increasing function in $\frac{\gamma}{2}$ when the size of the region in panel (b) in (**Figure 5**) is fixed. We set $\beta = \frac{\gamma}{2} \in [0, \frac{\pi}{4}]$ and use the polar coordinate transformations $(x', y') =$

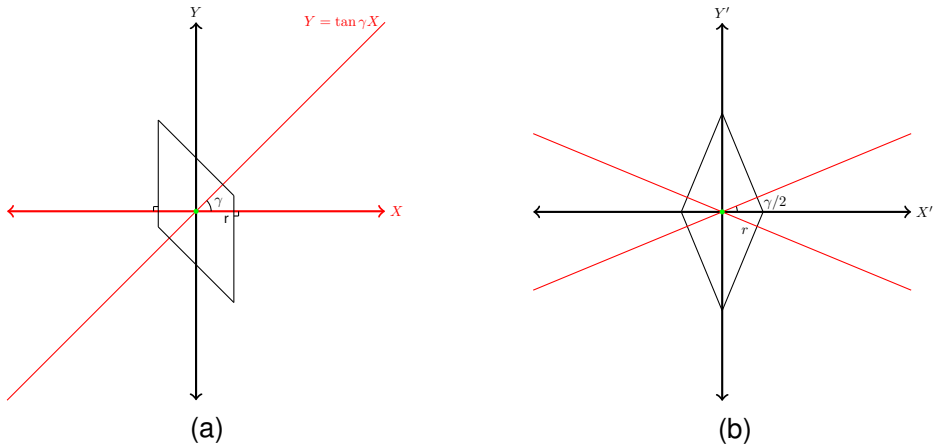


Figure 5: The diamond shaped region G_2 in (a) is the acceptance region for hypothesis 2 at the significance level $1-\mu(G_2)$. The diamond shaped region in panel (b) is obtained by rotating the region in panel (a) through the angle $\frac{\gamma}{2} \in [0, \frac{\pi}{4}]$ in a clockwise direction.

$(\rho\cos\theta, \rho\sin\theta)$. For a fixed significance level $\alpha \in [0,1]$, we get

$$\begin{aligned}
\frac{\pi}{2}(1-\alpha) &= \int_0^{\beta}\int_0^{\frac{r}{\cos(\beta-\theta)}} \exp\{-\frac{\rho^2}{2}\}\rho d\rho d\theta + \int_{\beta}^{\frac{\pi}{2}}\int_0^{\frac{r}{\cos(\beta-\theta)}} \exp\{-\frac{\rho^2}{2}\}\rho d\rho d\theta \\
&= \int_0^{\beta} 1-\exp\{-\frac{r^2}{2\cos^2(\beta-\theta)}\}d\theta + \int_{\beta}^{\frac{\pi}{2}} 1-\exp\{-\frac{r^2}{2\cos^2(\theta-\beta)}\}d\theta \\
&= \frac{\pi}{2} - \int_0^{\frac{\pi}{2}} \exp\{-\frac{r^2}{2\cos^2(\theta-\beta)}\}d\theta
\end{aligned}$$

We find $r' = \frac{dr}{d\beta}$ implicitly, by differentiating both side of the equation above with respect to β,

$$\begin{aligned}
0 &= -\frac{d}{d\beta}\int_0^{\frac{\pi}{2}} \exp\{-\frac{r^2}{2\cos^2(\beta-\theta)}\}d\theta \\
&= \int_0^{\frac{\pi}{2}} \exp\{-\frac{r^2}{2\cos^2(\beta-\theta)}\}\left[\frac{r}{\cos^2(\beta-\theta)}\right]r' d\theta \\
&\quad + \int_0^{\frac{\pi}{2}} \exp\{-\frac{r^2}{2\cos^2(\beta-\theta)}\}\left[\frac{r^2\sin(2(\beta-\theta))}{2\cos^4(\beta-\theta)}\right]d\theta.
\end{aligned}$$

to get

$$r' = -\frac{\int_0^{\frac{\pi}{2}} \exp\{-\frac{r^2}{2\cos^2(\beta-\theta)}\}\left[\frac{r^2\sin(2(\beta-\theta))}{2\cos^4(\beta-\theta)}\right]d\theta}{\int_0^{\frac{\pi}{2}} \exp\{-\frac{r^2}{2\cos^2(\beta-\theta)}\}\left[\frac{r}{\cos^2(\beta-\theta)}\right]d\theta} = \frac{\exp\{-\frac{r^2}{2\cos^2\beta}\} - \exp\{-\frac{r^2}{2\sin^2\beta}\}}{\int_0^{\frac{\pi}{2}} \exp\{-\frac{r^2}{2\cos^2(\beta-\theta)}\}\left[\frac{r}{\cos^2(\beta-\theta)}\right]d\theta}.$$

Since $\beta \in [0, \frac{\pi}{4}]$, we have $\cos^2\beta \geq \sin^2\beta$. As $\exp\{-\frac{r^2}{2*t}\}$ is a monotone increasing

function in t, we conclude that

$$\exp\{-\frac{r^2}{2\cos^2\beta}\} - \exp\{-\frac{r^2}{2\sin^2\beta}\} \geq 0,$$

where the equality holds only if $\beta = \frac{\pi}{4}$.

Since $\exp\{-\frac{r^2}{2\cos^2(\beta-\theta)}\}\left[\frac{r}{\cos^2(\beta-\theta)}\right] > 0$, we have

$$\int_0^{\frac{\pi}{2}} \exp\{-\frac{r^2}{2\cos^2(\beta-\theta)}\}\left[\frac{r}{\cos^2(\beta-\theta)}\right] d\theta > 0,$$

to conclude that

$$r' = \frac{\exp\{-\frac{r^2}{2\cos^2\beta}\} - \exp\{-\frac{r^2}{2\sin^2\beta}\}}{\int_0^{\frac{\pi}{2}} \exp\{-\frac{r^2}{2\cos^2(\beta-\theta)}\}\left[\frac{r}{\cos^2(\beta-\theta)}\right] d\theta} \geq 0.$$

Since the equality only holds when $\beta = \frac{\pi}{4}$, we can draw the conclusion that r is an increasing function for $\beta \in [0, \frac{\pi}{4})$.

B. Extensions to Asymptotic Bivariate Normal Distribution

The observation $(X_n, Y_n)^T$ is drawn from an asymptotic bivariate normal distribution with mean $\mu = (\mu_1, \mu_2)^T$ and positive definite variance Σ. We also assume that there exists a sequence of positive definite covariance matrices $\Sigma_n \xrightarrow{p} \Sigma$. We are interested in testing the null hypothesis $H_0 : \mu = (0, 0)^T$ versus the alternative $H_1 : \mu \in \mathcal{C}$, where $\mathcal{C} \in \mathbb{R}^2$. To simplify the problem, we apply the linear transformation $(X_n^*, Y_n^*)^T = \Sigma_n^{-\frac{1}{2}}(X_n, Y_n)^T$. Then $(X_n^*, Y_n^*)^T$ is from an asymptotic bivariate normal distribution with mean $\mu^* = \Sigma_n^{-\frac{1}{2}}\mu$ and identity variance. Equivalent to testing $H_0 : \mu = (0, 0)^T$ versus $H_1 : \mu \in \mathcal{C}$, we can test $H_0 : \mu^* = (0, 0)^T$ versus $H_1 : \mu^* \in \mathcal{C}^*$, where $\mathcal{C}^* \in \mathbb{R}$ where $\mathcal{C}_1^*$ and $\mathcal{C}_2^*$ are defined as

$$\mathcal{C}_1^* = \{(\mu_1^*, \mu_2^*)^T :\in \mathbb{R}^2 : (\mu_2^* - \mu_1^* \tan\gamma_1^*)(\mu_2^* - \mu_1^* \tan\gamma_2^*) \leq 0\},$$

and

$$\mathcal{C}_2^* = \{(\mu_1^*, \mu_2^*)^T :\in \mathbb{R}^2 : (\mu_2^* - \mu_1^* \tan\gamma_1^*)(\mu_2^* - \mu_1^* \tan\gamma_2^*) = 0\}.$$