Normal Spatial Model with Four Candidates in Three Dimensions:

Parameterization and Approximation

Chang Geun Song

(Research In-Progress)

Abstract

When there are two alternatives, the best way to decide one for group decision-making is majority rule. However, when there are more than two available options, it is not straightforward. Most of the previous literature of relevant fields focused on the three-candidate election, and when researchers were adopting spatial model, it is one-dimensional or two-dimensional. This paper targets the normal spatial model **(Good and Tideman, 1976)** with four candidates in three-dimensional attribute space. Most of our research is about the modeling process and technical details, parameterization of the models that can describe the actual rank data best. Furthermore, we test whether bimodal multivariate normal distribution assumption is more likely than unimodal multivariate normal distribution by AIC and BIC criteria.

## 1. Introduction

While diverse voting systems are introduced, researchers try to elect the most outstanding system. A lot of studies suggested variety of criteria to evaluate and verify voting systems. Many of those tests related the more than two-candidate cases and hence they required voters' ranking of the candidates or score of the candidates. The lack of rank or score data, it is difficult to be supported by empirical data, and it forced researchers to use statistical modeling that is assuming the distribution of preference ordering.

Widely used assumptions of the distribution is Impartial Culture (IC) and Impartial Anonymous Culture (IAC). IC assumes that each voter can have one of possible strict preference orderings with equal probability, and all possible profiles will have the same probability of realization. IAC assumes that one cannot differentiate the name of voters, hence profiles that consists of same elements are treated as one profile. Then assign equal probability on realization of each 'class' of profile.

Another popular modeling technique is spatial model. **Hotelling (1929)** and **Downs (1957)**, provided initial idea and clarified it. Main idea of the model called Hotelling-Downs is that voters ideal points distributed in a one-dimensional policy space, the voters prefer the candidate who has the closest political attribute to their ideal, and two or more candidates try to find their location to win the most votes. The model is used for game theory and public choice either. The spatial model is a generalized version of it: voters' ideal points and candidates' political attributes distributed in n-dimensional space.

This paper is meant neither as an introducing new theory, nor addressing actual empirical work such as test criteria or frequency of paradoxes. We mainly focus on the process that is the middle bridge: demonstrating modeling and its detail. We follow the normal spatial model **(Good and Tideman 1976)**. We assert significance of our work and differentiations from previous studies, in three aspects.

First, we assume the fourth candidate. Previous studies mostly consider three-candidate cases. Maybe presuming three candidates is more realistic, and be enough to analyze most of voting criteria or paradoxes including infamous Condorcet's paradox. However, that does not mean it is totally pointless to analyze four-candidate cases. **De Sinopoli, Dutta, & Laslier (2006)** assumed four-candidate election under approval voting system, and provided fairly engaging examples: the Condorcet winner receives NO votes, and strategic stability does not imply sincerity. **Niemi and Riker (1976)** noted that "And a third-place or even a fourth-candidate could exercise considerable influence over the choice of the president, as Henry Clay did in 1824 and George Wallace hoped to do in 1968." **Kurrild-Klitgaard, P. (2001)** brought up a possibility that the introduction of fourth candidate could make significant different outcome in an actual election, Prime ministerial candidates in the Danish election of 1994.

Second, we set the model with three dimensions. **Good and Tideman (1976)** provided all theoretical backgrounds of the normal spatial model for general n-candidates case. We do actualize the theory to the modeling process. As there is huge jump from one-dimensional space to two-dimensional space, dealing with three-dimensional space is a next level. We need to estimate much more parameters for location of candidates and voters' distributions, and calculating the probability density is another problem: With four candidates, in one-dimension setup, it is only calculating integral of cumulative density function of normal distribution six times. In three-dimension setup, one needs to calculate tri-variate normal distribution's density on twenty-four three dimensional shapes from a sphere cut by six planes.

We analyze large amount of synthetic elections.

(Capture the distribution of preferences, evaluating the voting rules, especially ranked voting, with many criterion)

## 2. Literature Review

### Spatial Model

After **Hotelling (1929), Downs (1957)** and **Good and Tideman (1976)**, it is widely used in literature. **Chamberlin and Cohen (1978), Merrill, S. (1984), Green-Armytage (2014), Green-Armytage, Tideman, and Cosman(2016)** assumed multiple candidates and multiple dimensions[1], and measure various criteria under different voting rules. Those studies simulated the model by random drawing of the candidates' and voters' locations from a multivariate normal distribution. On the other hand, we have additional pre-stage, which is estimating the best describing candidates' location and voters' distribution according to the real data.

### Simulation

**Plassmann and tideman (2012, 2014)** are more related to hour work.

[1] Chamberlin and Cohen (1978) assumed 4 candidates in 4 dimensions, and checked Condorcet criteria, Merrill, S. (1984) assumed 3,4,5,7 candidates in 1~4 dimensions,

Among 12 different model, NSM fits the data well.

Plassmann and Tideman (2014) evaluate different voting rules, check criteria, frequency of paradox…with NSM 3 cands in two d

- The normal spatial model with four candidates in three dimensions. Two-dimension setup is not enough to have 24 partitions that are corresponding to all possible strict preference ordering of four candidates (6 strict preference orderings are missing). Thorsten Matje (2016) showed invalidity of two-dimensional spatial model with four-candidate election in empirical framework.

Empirically it is not fit.

## 3. Analysis

## Pre-stage

Our modeling is about to estimating the frequency of the strict preference ordering. Hence, we define relevant presentations.

When we have three alternatives A, B, and C, and a voter prefer A to B, B to C, A to C then we will use ABC. If a voter prefers A to B and C, but is indifferent between B and C, then we will bind two indifferent options by parenthesis as A(BC). To present the number of voters who has preference as ABC, we use n[ABC].

## Data Process

We use parts of German Politbarometer data. The section, *Skalometer,* is ask respondents to independently evaluate a number of politicians by assigning one of 11 integers from -5 to 5. The greater integer means the respondent like more the politician. The survey began from March 1977 to December 2019. It is basically monthly data, but sometimes the respondents had different question set even in the same month. We consider them as independent surveys, and also divide the data into two by region (Western Germany and Eastern Germany) because of the same reason. 1,022 surveys are available after the process.

In each survey, 4~23 candidates are listed for questions and 179 to 3,345 voters participates. In average, 10.62 and 930.05 each.

First, we get rid of responses that are meaningless or careless to mitigate distortion from false preference reporting. Specifically, when a participant assigns one same integer to every candidate, or report with all empty we consider it as false report and eliminate his evaluation.

Based on score the respondent assigned, set the preference ordering. When the respondent did not assign a score, then we filled the score sampled from a multinomial distribution built by a frequency of scores the candidate wins from other voters. For convenience, we only count strict preference orderings and don't allow the respondent to be indifference between two or more candidates. If the participant assigns the same score on multiple candidates, we also set a multinomial distribution built by a frequency of others' preference orderings and sample from it. For instant, suppose three candidates A, B, and C, and if a respondent assigns 3, 3, 1 for each of them. Then the preference can be represented by (AB)C. Then, count the number of other voters who has preference ABC, and BAC, and the voters preference will be sampled from a multinomial distribution with p1 = n[ABC]/(n[ABC]+n[BAC]) and p2= n[BAC]/(n[ABC]+n[BAC]).

After disposing meaningless responses, filling the blanks, and breaking the indifference, we pick up top four candidates from the candidate pool. Here, “top four candidates” means four candidates top ranked by Condorcet method. i.e., four candidates would win the most in head-to-head election. When there exist cycle or ties that makes not possible to pick the top four candidates by Condorcet ranking, we apply Ranked Pair method (Tideman, 1987) or pick one that wins one-to-one election.

## Framework

For each survey, first, we pick top three candidates out of four, and we use the normal spatial model with three candidates in two-dimension, x-y plane; Finding estimated location of three candidates and bivariate normal distribution’s mod. Then we add z-axis, and expand the dimension, then find the best estimated location of four candidates and tri-variate normal distribution’s mod.

- Normal Spatial Model three-candidate two-dimensional (3C-2D)

With location of three candidates’ political attribute, partitioning the space by three bisect lines of possible pairs induces six area that are corresponding six strict preference order. Finding three points is mathematically identical to find the three bisect lines, it can be identified by two angles. We need two more parameters that can set the mod of the bivariate normal distribution of voters’ ideal points. By four parameters, the probability density on six area can be approximated, we set a multinomial distribution with p1, p2, p3, p4, p5, p6 that are corresponding to the probability density of each area. Four parameters are found by maximum likelihood method. We use Gauss-Newton algorithm, but the likelihood function does not have closed form and numerical differentiation is used for Jacobian.

- Normal Spatial Model four-candidate three-dimensional (4C-3D)

Add z-axis, we need one more parameter for setting the z-coordinate of the mod of the multivariate distribution. Three more parameters required for defining three planes that bisects D and each A, B, and C. From six plane, we have 24 partitions. Approximating the density on each partition, set a multinomial distribution with p1, p2, p3, …, p24, that are corresponding to the probability density of each three-dimensional partition. Eight parameters are found by maximum likelihood method. We also use Gauss-Newton algorithm and numerical differentiation. In Appendix 1 we will explain parameter estimation process in detail.

## Model Comparison ()

(We will compare

1. normal spatial model of 3C-2D and normal spatial model of 4C-3D

OR

2. normal spatial model of 4C-2D (from **Matje 2016**) and normal spatial model of 4C-3D

By comparing squared error or difference between parameters of multinomial distribution to actual share of strict preference orderings.

*expecting result could be showing 4C-3D model is as good as 3C-2D at least, and or much better than 4C-2D)

## Extended model

Assume that the distribution of voters' ideal points as bimodal normal distribution. Not estimating from the scratch, we assume two mods that are near around the mod that found with unimodal assumption. The idea is that the share of the 24 strict preference order is well estimated by unimodal assumption, and we want the new distribution with bimodal assumption but not very different so that the estimation of share of 24 strict preference order changes not very much.

We have a spherical multivariate normal distribution in three-dimensional space, make two mods symmetrically departed from the old mod, then we have 13 ways.

(x+.x-)(z+,z-) (y+,y-)

(x+y+,x-y-)(x+y-,x-y+) (x+z+,x-z-)(x+z-,x-z+)(y+z+,y-z-)(y+z-,y+z-)

(x+y+z+,x-y-z-)(x+y+z-,x-y-z+)(x+y-z+,x-y+z-)(x-y+z+,x+y-z-)

### How much should they far apart ()

When it comes to taking bimodal assumption, how should two mods be far not to accept unimodal normal distribution hypothesis?

### Selection of Model

Compare the Unimodal assumption and Bimodal assumption by AIC and BIC.

## 4. Result

One criticism about Plurality rule is that it constrains the third options.

- diversification

- This paper provides a milestone to known as exist but the unexplored.

- By checking the changes in likelihood ratio, AIC, and BIC we can have partial hint to discussion about bimodal assumption.

- eight variables, lots of approximations, and numerical method

>>>>> room to be more accurate, and time-efficient.

1. mathematically; better approximation and optimization method.

2. coding skills.-> time-saving -> iterate more -> better approximation

**References**

[] Downs A. (1957). An Economic Theory of Democracy. Harper & Row, New York.

[] Niemi, R. G., & Riker, W. H. (1976). The Choice of Voting Systems. Scientific American, 234(6), 21–27. http://www.jstor.org/stable/24950368

[] Good, I. J. & Tideman T. N. (1976). From individual to collective ordering through multidimensional attribute space. Proceedings of the Royal Society of London. A. Mathematical and Physical Sciences, 347(1650):371–385.

[] Hotelling H. (1929). Stability in competition. Econ. J. 39, 41–57.

[] Ornstein, J. T. & Norman, R. Z. (2014). Frequency of monotonicity failure under Instant Runoff Voting: estimates based on a spatial model of elections. Public Choice 161, 1–9.

[] Tideman T.N., & Plassmann F. (2012). Modeling the Outcomes of Vote-Casting in Actual Elections. In: Felsenthal D., Machover M. (eds) Electoral Systems. Studies in Choice and Welfare. Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-642-20441-8_9

[] Plassmann, F., & Tideman, T. N. (2014). How frequently do different voting rules encounter voting paradoxes in three-candidate elections? Social Choice and Welfare, 42(1), 31–75.

[] Kurrild-Klitgaard, P. (2001). An Empirical Example of the Condorcet Paradox of Voting in a Large Electorate. Public Choice 107, 135–145 https://doi.org/10.1023/A:1010304729545

[] Green-Armytage, J., Tideman, T. N., & Cosman, R. (2016). Statistical evaluation of voting rules. Social Choice and Welfare, 46(1), 183-212.

[] Merrill, S. (1984). A Comparison of Efficiency of Multicandidate Electoral Systems. American

Journal of Political Science, 28(1), 23–48. https://doi.org/10.2307/2110786

[] Chamberlin, J. R., & Cohen, M. D. (1978). Toward applicable social choice theory: A comparison of social choice functions under spatial model assumptions. American Political Science Review, 72(4), 1341-1356.

[] Green-Armytage, J. (2014). Strategic voting and nomination. Social Choice and Welfare, 42(1), 111-138.

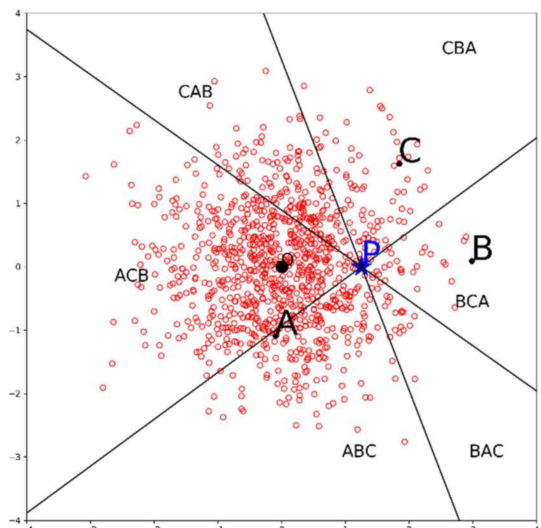
[] De Sinopoli, F., Dutta, B. & Laslier, JF. Approval voting: three examples. Int J Game Theory 35, 27–38 (2006). https://doi.org/10.1007/s00182-006-0053-2

Appendix. Parameter Estimation and Modeling in detail

- NSM, 3C-2D

[objective]

Finding the best guess of location of three candidates and the voters' distribution on the two-dimensional attribute space.



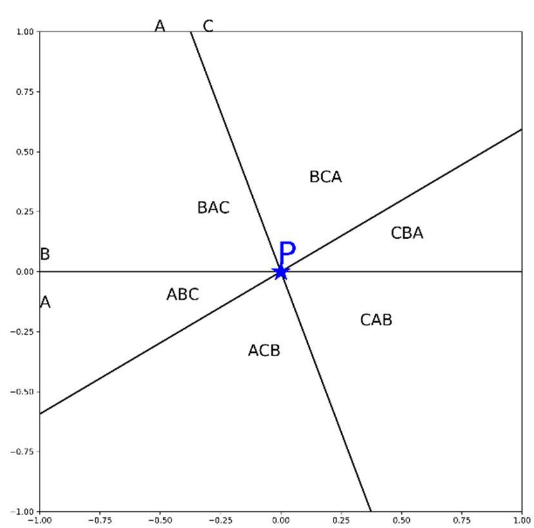

Given three points and normal distribution, there are three lines that bisects from each pair of points: A and B, B and C, A and C. Three bisects make 6 partitions, and each area corresponding to strict preference orderings. The density on each area are estimated share of each strict preference ordering.

Finding three bisects is equivalent to finding three points. When we rotate three lines and the mod of normal distribution against the point P, intersection of three lines, induces same density for six areas.

For convenience, set the origin at the P and fix a line that is bisecting A and B (A|B) at x-axis. Then we can define two bisect lines B|C, A|C with the angle between A|B and B|C (alpha), and between B|C and A|C (beta). We also need two more parameters to set the mod of the normal distribution (x,y).

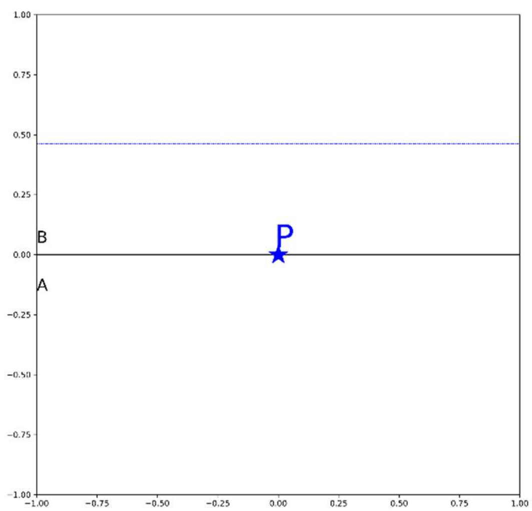
[step1. Initial guess]

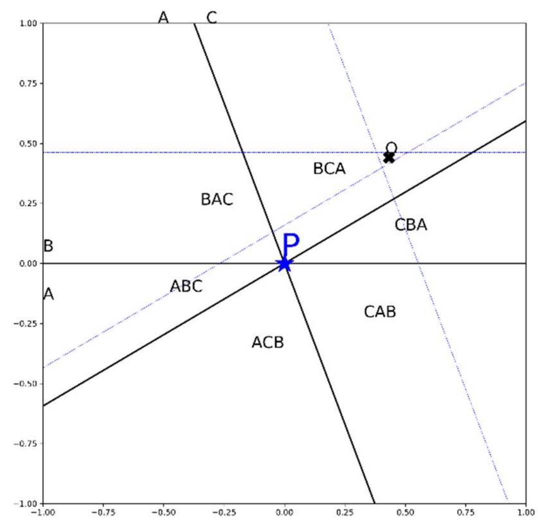
Partition ABC-CBA, ACB-BCA, BAC-CAB, located oppositely, and the vertically opposite angle should be the same. Summing up the number of voters that have ABC and CBA, and set the angle proportional to the share of the voters. Theta_ABC + theta_CBA = 360 * (n[ABC]+n[CBA]/total number of voters). Because, if the mod of the normal distribution at P, the density of each partition has one-to-one corresponds to the angle. Get other four angles, then we have alpha_0 and beta_0.




Each bisects, divide the space into two areas, two areas should provide the share of each candidate when the mod of normal distribution is on the line <figure>. Do same thing for other two lines.

We have a triangular part from newly constructed three parallel lines. Pick a center of the gravity (for convenient calculation) as an initial mod of the normal distribution <figure>.



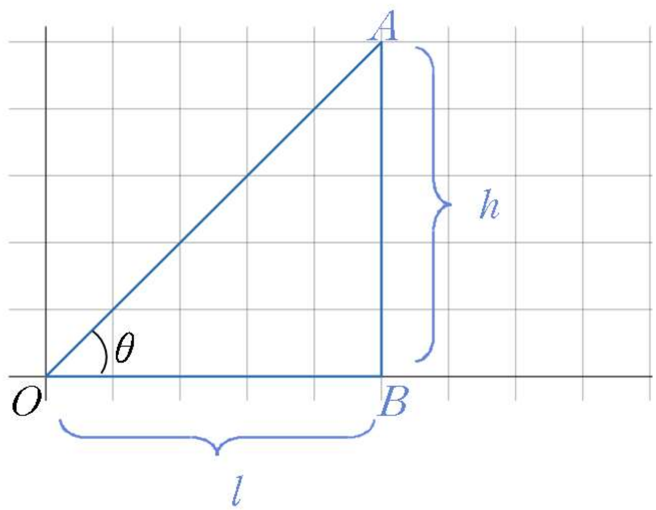




Given example, the mod of normal distribution located BCA area

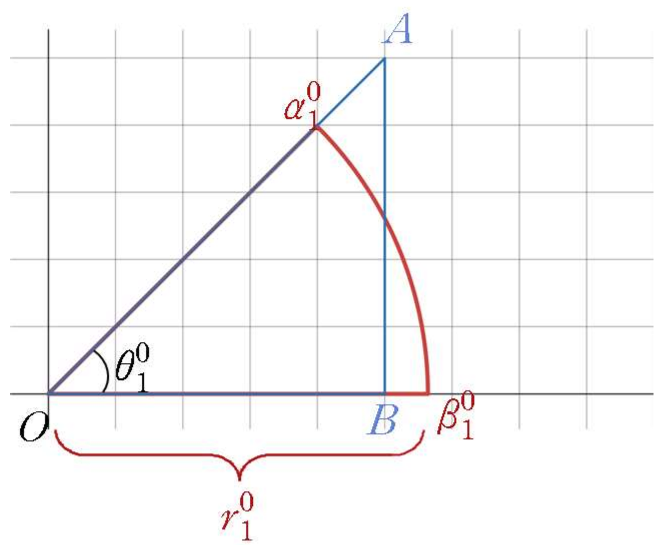
[step2. density calculation]-groundwork

Direct way of calculating the density of each partition is not available. We are starting from approximating the density within a right triangle with the mod of the bivariate normal distribution.

Suppose a triangle OAB which has l for length, h for height. O is the mod of bivariate normal distribution, and angle O is $\theta$ degree angle <figure>.




Since calculating the density of a sector form is easy, we calculate the density of a sector form Oαβ which has the same area with the triangle, and use it as an approximation of the density of the right triangle OAB <figure>.




Assume the bivariate normal distribution is standard normal with 0 covariance or,

$$\boldsymbol{\mu} = [0,0],\ \boldsymbol{\Sigma} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

Then the density of the sector with angle $\theta$ in radian and radius r is as follows

$$\Pr(R \leq r; \theta) = \frac{\theta}{2\pi}\left(1 - exp\left(-\frac{r^2}{2}\right)\right)$$

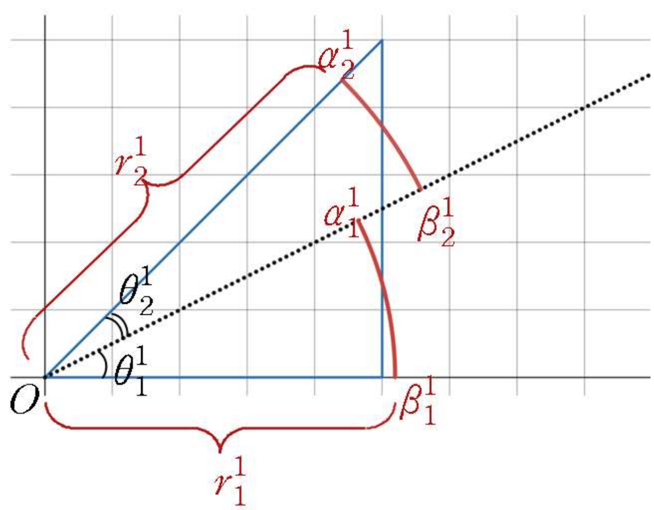
The radius that makes the area of the triangle and the sector equal can be found by $r^2 = \frac{lh}{\theta}$,

hence we have alternative representation, the function of l and h.

$$Pr(R \leq r; \theta) = P(l, h) = \frac{\theta}{2\pi}\left(1 - exp\left(-\frac{1}{2}\frac{lh}{\theta}\right)\right)$$

where $\theta = \tan^{-1}\left(\frac{l}{h}\right)$.

To get the finer approximation, divide the triangle into two to let the height equal h/2. Then we use density of each sector that has equal area to each triangle.

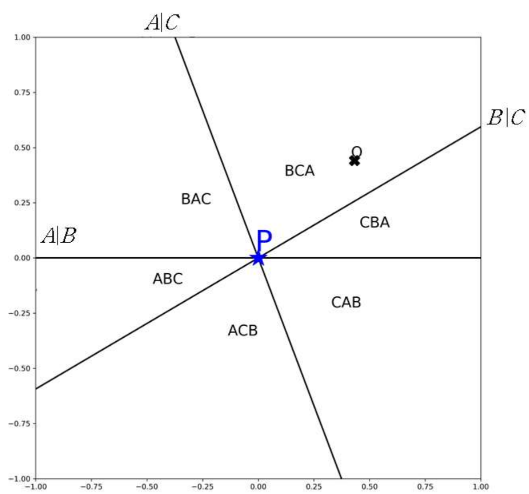



Approximating the density of triangle with N numbers of sector forms is

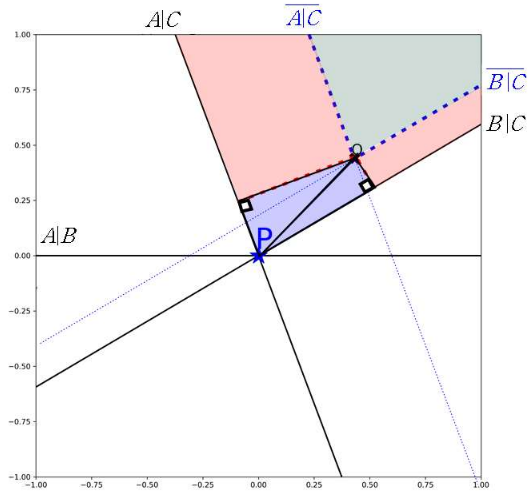
$$P(l,h) = \sum_{n=1}^{N}\left(\frac{\theta_n}{2\pi}\left(1 - exp\left(-\frac{1}{2N}\frac{lh}{\theta}\right)\right)\right)$$

where $\theta_n = \tan^{-1}\left(\frac{n}{N}\frac{l}{h}\right) - \tan^{-1}\left(\frac{n-1}{N}\frac{l}{h}\right)$.

[step2. density calculation]

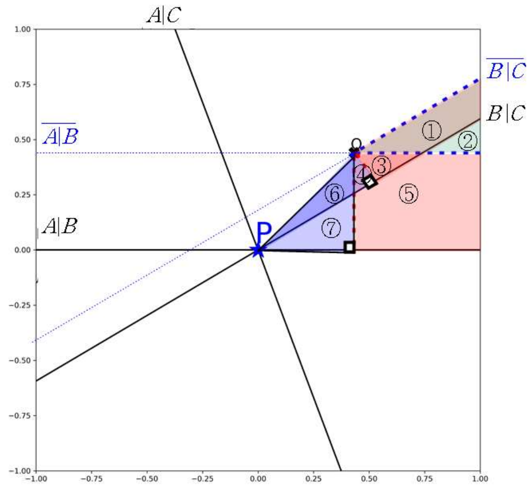



1. the partition that has the mod of normal distribution (partition-BCA)




The partition that has the mod of normal distribution can be divided into a sector with infinite radius (green), two half-infinite bar (red), two right triangles (blue). The density of a sector with infinite radius equals a ratio the angle between A|C and B|C to $2\pi$. For two half-infinite bars, the density can be found by density function of normal distribution with the width of the bar. Using circular sector approximation, we get the density of triangles. By adding all, one could get the density of partition-BCA.

2. the partition next to the partition that has the mod (partition-CBA)

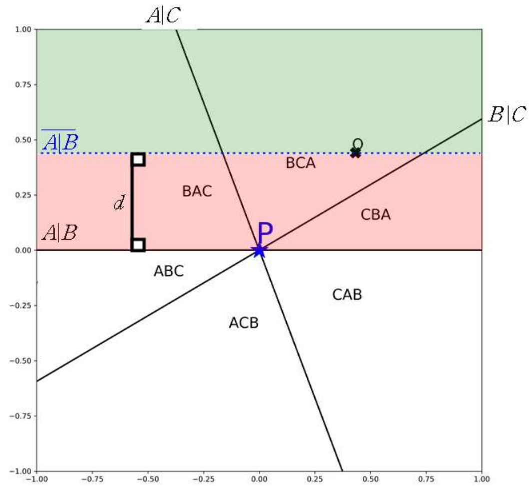



Graphically, the partition is ②+⑤+⑦. We calculate one circular sector with infinite radius (①+②), two half-infinite bar, (①+③, ③+④+⑤), two right triangles (④+⑥, ⑥+⑦), and combine them.

(①+②)+(③+④+⑤)- (①+③)-(④+⑥)+(⑥+⑦) = ②+⑤+⑦

3. Remaining partitions

From density of two partitions, we can calculate other partitions. For instance, to calculate the density of partition-BAC, we subtract the density of above two (partition-BCA and partition CBA) from the density of space above the x-axis. Which equals $\Phi(\mathrm{d})$ where $\Phi(\cdot)$ is the cumulative density function of standard normal distribution.




.

Likewise, all other partitions (that of ABC, ACB, CAB) can be derived. For six densities for six partitions, set the initial parameters of multinomial distribution $(p_1, p_2, p_3, p_4, p_5, p_6)$.

[step3. maximizing LLH]

With $(p_1, p_2, p_3, p_4, p_5, p_6)$, probability mass function of multinomial distribution is

$$f(p_1, p_2, p_3, p_4, p_5, p_6) = \frac{Q!}{q_1! \cdots q_6!} \prod_{i=1}^{6} p_i^{q_i}$$

where $(q_1, q_2, q_3, q_4, q_5, q_6)$ is number of each event happens and $\sum_{i=1}^{6} q_i = Q$.

the log likelihood function is, of multinomial distribution

$$llh(p_1, \cdots, p_6; q_1, \cdots, q_6) = ln(Q!) - \sum_{i=1}^{6} ln(q_i!) + \sum_{i=1}^{6} q_i ln(p_i)$$

In our model, we have four parameters $(\mathrm{x}, \mathrm{y}, \alpha, \beta)$ decides $(\mathrm{p}_1, p_2, p_3, p_4, p_5, p_6)$. For given $(\mathrm{q}_1, \mathrm{q}_2, \mathrm{q}_3, \mathrm{q}_4, \mathrm{q}_5, \mathrm{q}_6)$, $ln(Q!)$ and $\sum_{i=1}^{6} ln(q_i!)$ are constants and maximization problem is

$$\max_{\mathrm{x}, \mathrm{y}, \alpha, \beta} \mathrm{L}(\mathrm{x}, \mathrm{y}, \alpha, \beta; q_1, \cdots, q_6) = \sum_{i=1}^{6} q_i ln(p_i(\mathrm{x}, \mathrm{y}, \alpha, \beta))$$

We don't have closed form of function that induces $(\mathrm{p}_1, p_2, p_3, p_4, p_5, p_6)$ from $(\mathrm{x}, \mathrm{y}, \alpha, \beta)$, hence we cannot use a first order or second order method properly, and we use numerical differentiation and use approximation of them. For this paper, we use gauss-newton algorithm.

- NSM, 4C-3D

[objective]

Finding the best guess of location of four candidates and the voters' distribution on the three-dimensional attribute space.

[step1. Parameters, vectors and plane]

[step2. density calculation, tiling a sphere]

For a multivariate normal distribution with $(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ where $\boldsymbol{\mu}$ is px1 mean vector, and $\boldsymbol{\Sigma}$ is pxp variance-covariance matrix,

$$(\mathbf{x} - \boldsymbol{\mu})^{\mathbf{T}} \boldsymbol{\Sigma}^{-1} (\mathbf{x} - \boldsymbol{\mu}) \sim \chi_{\mathbf{p}}^{2}$$

Our model assumes standard trivariate normal distribution,

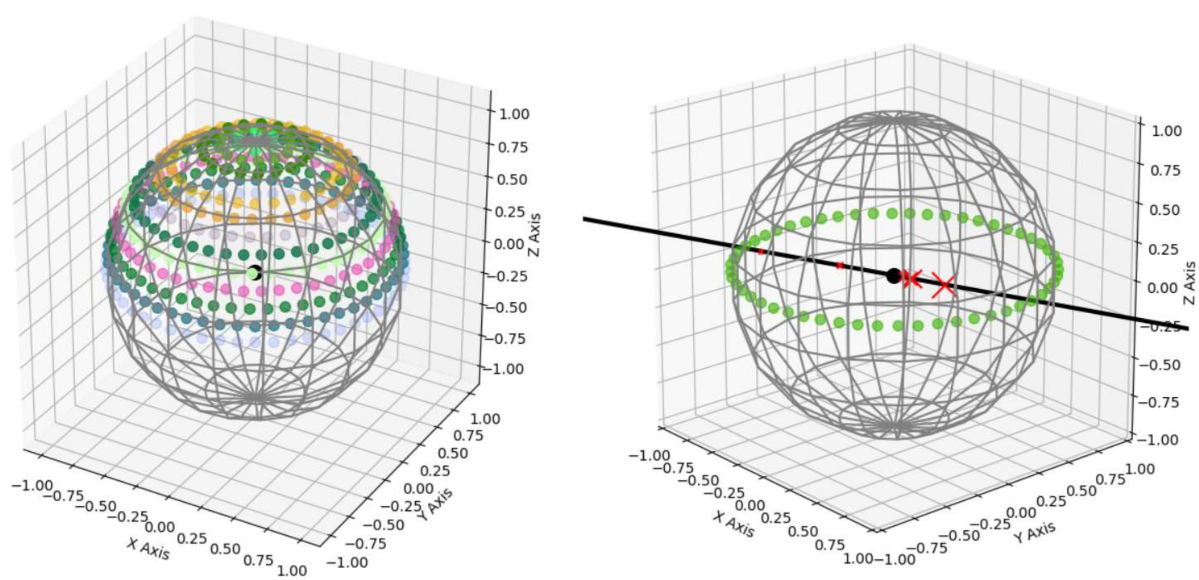
$$\boldsymbol{x^2 + y^2 + z^2 \sim \chi_3^2}$$

For a given radius r, the equation of the sphere with the center of sphere as O is

$$\boldsymbol{x^2 + y^2 + z^2 = r^2}$$

Therefore, the density of within-sphere, $\Pr[\boldsymbol{x^2 + y^2 + z^2 \leq r^2}]$, is $\Pr[\boldsymbol{\chi_3^2 \leq r^2}]$.

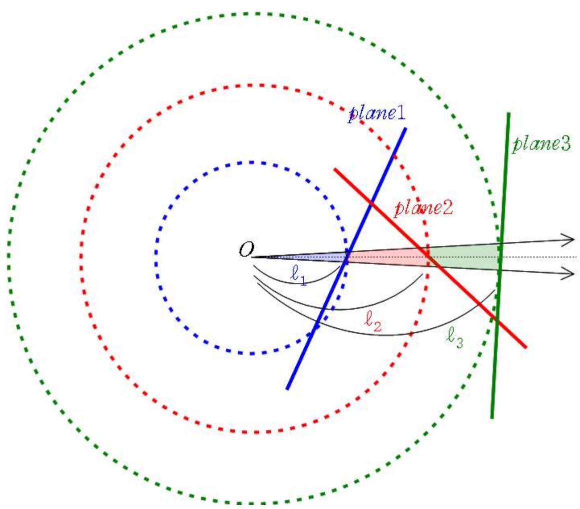
Consider a sphere that has the center of the sphere at O, the mod of the trivariate standard normal distribution. Put N number of points equally distributed on the surface of the sphere. Consider identical shape of tiles that cover the whole sphere, and has each point as the center of each tile.




For a given tile, imagine a three-dimensional structure, an infinitely tall pyramid, that made of lines originate from the O and pass through each sides of a tile. Some of six planes might cut the pyramid, and each partition will be a part of some strict preference ordering.

The idea is approximating each cut part of the pyramid with 1/N part of spheres with different radius. Consider a straight line that start from O, passes through the center of the tile. Let the distance from O to where the straight line meets the first cut edge as $l_1$. If there exist another cut then let the distance from O to where the straight line meets the second edge as $l_2$, and so forth.

We approximate the density of the first chunk as $\frac{1}{N}\Pr[\boldsymbol{\chi_3^2} \leq \mathbf{l_1}^2]$. For the second chunk, it is $\frac{1}{N}(\Pr[\boldsymbol{\chi_3^2} \leq \mathbf{l_2}^2] - \Pr[\boldsymbol{\chi_3^2} \leq \mathbf{l_1}^2])$ and $\frac{1}{N}(\Pr[\boldsymbol{\chi_3^2} \leq \mathbf{l_3}^2] - \Pr[\boldsymbol{\chi_3^2} \leq \mathbf{l_2}^2])$ for the third. When there is no more cut, then for the remaining giant chunk, the density is approximated to $\frac{1}{N}(1 - \Pr[\boldsymbol{\chi_3^2} \leq \mathbf{l_3}^2])$.




We divided the sphere by 2,000 tiles, so with 2,000 straight lines. For each of lines, we check the intersection with six-plane and approximate the density of chunks, define which strict preference ordering is corresponding, and summing up by the preference. It ends up with 24 numbers.

[step3.maximizing LLH]

In our model, we have eight parameters that deciding $(\mathrm{p}_1, p_2, \cdots, p_{24})$ . For given $(\mathrm{q}_1, \mathrm{q}_2, \mathrm{q}_3, \mathrm{q}_4, \mathrm{q}_5, \mathrm{q}_6)$, $ln(Q!)$ and $\sum_{i=1}^{6} ln(q_i!)$ are constants.

$$\underset{\mathrm{x, y, \alpha, \beta, \gamma, \delta, \varepsilon}}{max} \mathrm{L}(\mathrm{x, y, z, \alpha, \beta, \gamma, \delta, \varepsilon}; q_1, \cdots, q_{24}) = \sum_{i=1}^{24} q_i ln(p_i(\mathrm{x, y, z, \alpha, \beta, \gamma}))$$

Have exactly same problem with previous parameter estimation in 3C-2D, we use gauss-newton algorithm with numerical differentiation.