Journal of Xue Gong for Fall semester 2015

November 17, 2015

1 Week of September 8 - 14

I am working on finding the equilibrium points to the gradient flow outside of the symmetric set when using a rank 2 tensor to approximate a rank 2 target.

When setting the partial derivatives of the error function to be 0, the equations are very complicated. However, I found two necessary conditions for the equilibrium points.

$$\frac{n_j(\vec{\alpha})}{p_j(\vec{\alpha}-\vec{\beta})} = \frac{n_k(\vec{\alpha})}{p_k(\vec{\alpha}-\vec{\beta})}, \text{ and } \frac{n_j(\vec{\beta})}{p_j(\vec{\alpha}-\vec{\beta})} = \frac{n_k(\vec{\beta})}{p_k(\vec{\alpha}-\vec{\beta})}.$$

When consider the case where $\alpha_1 \neq \alpha_2 = \cdots = \alpha_d = \alpha$ and $\beta_1 = \cdots = \beta_d = \beta$, I found some results regarding α_1, α, and β at the equilibrium points.

2 Week of September 15 - 21

After discussing with Dr. Mohlenkamp, we found that for the case $\alpha_1 \neq \alpha_2 = \cdots = \alpha_d = \alpha$ and $\beta_1 = \cdots = \beta_d = \beta$, the equilibrium point has to satisfy $\alpha_1 = \alpha$. Thus, this situation does not exist.

Then I start to study the case where $\alpha_1 \neq \alpha_2 = \cdots = \alpha_d = \alpha$ and $\beta_1 \neq \beta_2 = \cdots = \beta_d = \beta$. Using the equations given from last week's journal and after simplifying the expressions by some trigonometric identities, we have

$$\begin{aligned} 0 = &\cos(\alpha-\beta)\sin(\alpha_1-\beta_1)\sin(\alpha_1-\alpha)\{(\cos\alpha)^{d-2} + z[\cos(\alpha-\phi)]^{d-2}\} \\ &+ \sin(\alpha-\alpha_1-\beta+\beta_1)\{(\cos\alpha)^{d-1}\sin\alpha_1 + z[\cos(\alpha-\phi)]^{d-1}\sin(\alpha_1-\phi)\}, \end{aligned} \tag{1}$$

and

$$\begin{aligned} 0 = &\cos(\alpha-\beta)\sin(\alpha_1-\beta_1)\sin(\beta_1-\beta)\{(\cos\beta)^{d-2} + z[\cos(\beta-\phi)]^{d-2}\} \\ &+ \sin(\alpha-\alpha_1-\beta+\beta_1)\{(\cos\beta)^{d-1}\sin\beta_1 + z[\cos(\beta-\phi)]^{d-1}\sin(\beta_1-\phi)\}. \end{aligned} \tag{2}$$

For $\alpha \neq \beta$ and $\alpha_1 \neq \beta_1$, one possible solution for these two equations can be found by simply setting the expressions in all the brackets be 0. Then we can obtain for $\phi \neq 0$ and some integer k,

$$\alpha_1 + \alpha = \phi - \frac{\pi}{2} + k\pi, \text{ and } \beta_1 + \beta = \phi - \frac{\pi}{2} + k\pi.$$

These equations are the same as what we found in the G1T2 case. We might be able to use these two equations to find out the equilibrium points looking at the partial derivatives for the error functions again. We can also try to find other solutions for (1) and (2), but I tried several times and did not find out how to do this yet.

If we let $\beta = \phi - \alpha$ and $\beta_1 = \phi - \alpha_1$, then we can see that the LHS of (1) = the LHS of (2) when $\cos\alpha = \cos(\alpha - \phi)$. If $\phi \neq 0$, then $\alpha = \phi/2 + k\pi/2$. So when $\alpha \to \phi/2$ and $\alpha_1 \to \phi/2 \pm \pi/2$, the LHS of both (1) and (2) are approaching 0. This agrees with the numerical results from Dr. Mohlenkamp.

3 Week of September 22 - 28

This week I have been preparing for my presentation on the stability of non-symmetric equilibrium points for the G1T2 case. I also checked some calculations I did to make sure all the results are correct.

4 Week of September 29 - October 5

For the case where $\alpha_1 \neq \alpha_2 = \cdots = \alpha_d = \alpha$ and $\beta_1 \neq \beta_2 = \cdots = \beta_d = \beta$, I have found three solutions to equations (1) and (2).

1. The first trivial solution is $\alpha_1 = \beta_1$ and $\alpha = \beta$. Then we can simplify

$$-\frac{\partial}{\partial \alpha_j} E_\lambda(G_2) = \frac{2n(\vec{\alpha})n_j(\vec{\alpha})}{\lambda + 2} = 0, \text{ for all } j.$$

   - One equilibrium point is when $n(\vec{\alpha}) = 0$. This again gives us the maximum point.
   - Another solution is when $n_j(\vec{\alpha}) = n_j(\vec{\beta}) = 0$ for all j. Then we have a system of two equations about two variables. This turns out to have the solutions when

$$\alpha_1 + \alpha = \phi - \frac{\pi}{2} + k\pi, \text{ and } \beta_1 + \beta = \phi - \frac{\pi}{2} + k\pi.$$

2. Another solution is $\alpha - \beta = \pm\pi/2$ and $\alpha_1 - \beta_1 = \pm\pi/2$. We can obtain that

$$-\frac{\partial}{\partial \alpha_j} E_\lambda(G_2) = \frac{2n(\vec{\alpha})n_j(\vec{\alpha})}{\lambda + 1} = 0, \text{ for all } j.$$

   The solutions to $n_j(\vec{\alpha}) = n_j(\vec{\beta}) = 0$ need to satisfy

$$\frac{\tan(\alpha_1 - \phi)}{\tan\alpha_1} = \frac{\tan(\alpha - \phi)}{\tan\alpha} \text{ and } \frac{\tan(\alpha_1 - \phi)}{\tan\alpha_1} = \left(\frac{\tan(\alpha - \phi)}{\tan\alpha}\right)^{d-1}.$$

   This seems to be true when $d = 2$, or d is even and $\frac{\tan(\alpha-\phi)}{\tan\alpha} = -1$. We can solve for α and α_1, then find β and β_1.

3. $\alpha_1 + \alpha = \phi - \frac{\pi}{2} + k\pi$, and $\beta_1 + \beta = \phi - \frac{\pi}{2} + k\pi$. In this case, for the blue brackets in (1) and (2), we can find the solution:

$$\alpha \text{ (and } \beta) = \arctan\left(\frac{(-z)^{-1/(d-2)} - \cos\phi}{\sin\phi}\right).$$

   - When d is odd, there is only one solution in $(-\pi/2, \pi/2)$, then $\alpha = \beta$, and $\alpha_1 = \beta_1 = \phi - \alpha + \pi/2 + k\pi$.
   - When d is even and $z < 0$, there are two solutions in $(-\pi/2, \pi/2)$. So we can have $\alpha \neq \beta$ and $\alpha_1 \neq \beta_1$.
   - When d is even and $z > 0$, there is no solution. So the equilibrium point in this form does not exist.

5 Week of October 6 - 12

Ignoring the solutions where $\alpha_1 = \beta_1$ and $\alpha = \beta$ because this becomes the rank 1 case, we first consider the special solution in case 2 from last week. When d is even and $\frac{\tan(\alpha-\phi)}{\tan\alpha} = -1$, the solutions are

$$\alpha_1 = \beta_2 = \cdots = \beta_d = \beta = \phi/2 \text{ and } \beta_1 = \alpha_2 = \cdots = \alpha_d = \phi/2 - \pi/2;$$

$$\text{or } \alpha_1 = \beta_2 = \cdots = \beta_d = \beta = \phi/2 - \pi/2 \text{ and } \beta_1 = \alpha_2 = \cdots = \alpha_d = \phi/2.$$

These two solutions are equivalent by interchanging α's and β's, thus we only consider the first solution. It is easy to find that

$$\frac{\partial^2}{\partial\alpha_j\partial\alpha_k}E_\lambda(G_2) = \frac{2}{1+\lambda}n(\vec{\alpha})n_{jk}(\vec{\alpha}), \text{ and } \frac{\partial^2}{\partial\beta_j\partial\beta_k}E_\lambda(G_2) = \frac{2}{1+\lambda}n(\vec{\beta})n_{jk}(\vec{\beta}).$$

This case has similar results as the T2G1 case. After calculations, we can find that the Hessian matrix is

$$H = \frac{2}{1+\lambda}\begin{bmatrix} H_A & \mathbf{0} \\ \mathbf{0} & H_B \end{bmatrix},$$

where

$$H_A = \begin{bmatrix} n^2(\vec{\alpha}) & n^2(\vec{\alpha}) & \cdots & \cdots & n^2(\vec{\alpha}) \\ n^2(\vec{\alpha}) & n^2(\vec{\alpha}) & A & \cdots & A \\ \vdots & A & \ddots & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & A \\ n^2(\vec{\alpha}) & A & \cdots & A & n^2(\vec{\alpha}) \end{bmatrix}, \text{ and } H_B = \begin{bmatrix} n^2(\vec{\beta}) & n^2(\vec{\beta}) & \cdots & \cdots & n^2(\vec{\beta}) \\ n^2(\vec{\beta}) & n^2(\vec{\beta}) & B & \cdots & B \\ \vdots & B & \ddots & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & B \\ n^2(\vec{\beta}) & B & \cdots & B & n^2(\vec{\beta}) \end{bmatrix}.$$

We can use the householder reflector $\tilde{P} = \begin{bmatrix} P & \mathbf{0} \\ \mathbf{0} & P \end{bmatrix}$ where P is the householder reflector in the T2G1 case and obtain that

$$PHP = \frac{2}{1+\lambda}\begin{bmatrix} \Omega_A & \mathbf{0} \\ \mathbf{0} & \Omega_B \end{bmatrix},$$

where

$$\Omega_A = \begin{bmatrix} n^2(\vec{\alpha}) & \sqrt{d-1}n^2(\vec{\alpha}) & & & \\ \sqrt{d-1}n^2(\vec{\alpha}) & n^2(\vec{\alpha}) + (d-2)A & & & \\ & & n^2(\vec{\alpha}) - A & & \\ & & & \ddots & \\ & & & & n^2(\vec{\alpha}) - A \end{bmatrix},$$

and Ω_B can be obtained by replacing α with β and replacing A with B in Ω_A. In fact, we can calculate that

$$\begin{aligned} n^2(\vec{\alpha}) - A &= (1-z)^2 \cos^2(\phi/2) \sin^{2d-4}(\phi/2); \\ n^2(\vec{\beta}) - B &= (1-z)^2 \sin^2(\phi/2) \cos^{2d-4}(\phi/2). \end{aligned}$$

Thus, $n^2(\vec{\alpha}) - A \geq 0$ and $n^2(\vec{\beta}) - B \geq 0$. Therefore, we can conclude that when $z \neq 1$ and $\phi \neq 0$, there are $2d-2$ positive eigenvalues and 2 negative eigenvalues.

6 Week of October 13 - 19

6.1 G2T2 case

Consider case 3 in section 4, when d is even and $z < 0$, we have the solution

$$\begin{aligned} \alpha &= \arctan\left(\frac{(-z)^{-1/(d-2)} - \cos\phi}{\sin\phi}\right), \quad \alpha_1 = \phi - \alpha - \pi/2 + k_1\pi; \\ \beta &= \arctan\left(\frac{-(-z)^{-1/(d-2)} - \cos\phi}{\sin\phi}\right), \quad \beta_1 = \phi - \beta - \pi/2 + k_2\pi. \end{aligned}$$

In this case, we can find that $n_j(\vec{\alpha}) = n_j(\vec{\beta}) = 0$ for all j. Then plug this solution into the expression for $\frac{\partial}{\partial \alpha_j} E_\lambda(G_2)$, we have

$$\frac{\partial}{\partial \alpha_j} E_\lambda(G_2) = -2p_j(\vec{\alpha} - \vec{\beta}) \frac{n(\vec{\alpha})n(\vec{\beta})[(1+\lambda)^2 + p^2(\vec{\alpha} - \vec{\beta})] - (1+\lambda)(n^2(\vec{\alpha}) + n^2(\vec{\beta}))p(\vec{\alpha} - \vec{\beta})}{[(1+\lambda)^2 - p^2(\vec{\alpha} - \vec{\beta})]^2}.$$

If $z \neq -1$ ($z = -1$ leads to the solution in section 5), we can check that $\alpha - \beta \neq \pm\pi/2$, then $p_j(\vec{\alpha} - \vec{\beta}) \neq 0$. Therefore, the equilibrium point to the gradient flow needs to satisfy the numerator (blue part) is 0. This can be simplified and treated as a function of $1 + \lambda$:

$$(1+\lambda)^2 - \cos^{d-1}(\alpha - \beta)\cos(\alpha_1 - \beta_1)\left[\left(\frac{\cos\alpha}{\cos\beta}\right)^{d-2} + \left(\frac{\cos\beta}{\cos\alpha}\right)^{d-2}\right](1+\lambda) + \cos^{2d}(\alpha - \beta) = 0.$$

So only when λ is the positive solution to the above equation, the listed $(\alpha, \alpha_1, \beta, \beta_1)$ is a critical point.

6.2 G1T2 case

In the case when $\alpha_1 \neq \alpha_2 = \cdots = \alpha_d$, we already know the eigenvalues to the Hessian matrix are

$$x_{1,2} = \frac{1}{1+\lambda}\left[(d-2)A + 2n^2 \pm \sqrt{(d-2)^2A^2 + 4(d-1)n^4}\right],$$
$$x_3 = \cdots = x_d = \frac{2}{1+\lambda}\left(n^2 - A\right).$$

When $n^2 - A > 0$, we have $x_1 > 0, x_3 = \cdots = x_d > 0$ and $x_2 < 0$.

- After some calculations, we can find that when $0 < A < n^2$,
$$x_2 < 0 < x_3 = \cdots = x_d < x_1, \text{ and } |x_2| < x_1.$$
and
$$-\frac{x_2}{x_3} \leq \frac{\sqrt{d-1}-1}{1-\cot\alpha_1\cot(\alpha_1-\phi)}.$$
Thus, this ratio can be either > 1 or < 1 depending on the parameters.

- When $A < 0$, we have
$$\frac{x_1}{x_3} \leq \frac{\sqrt{d-1}+1}{1-\cot\alpha_1\cot(\alpha_1-\phi)}.$$
Thus, it depends on the parameters to determine whether $x_1 < x_3$ or $x_1 > x_3$, but in this situation, we know that $-\frac{x_2}{x_3} > 1$, i.e., $|x_2| > x_3$.

7 Week of October 20 - 26

Continue with last week's work in section 6.2, for G1T2 case when $\alpha_1 \neq \alpha_2 = \cdots = \alpha_d$,

- When $0 < A < n^2$, we have $x_2 < 0 < x_3 = \cdots = x_d < x_1$. If we want $-\frac{x_2}{x_3} < 1$, then the condition can be simplified to
$$0 < \tan\alpha_1\tan(\alpha_1-\phi) < \frac{d-3}{d-5}.$$
We can solve these inequalities to obtain
$$z < 0, (-z)^{1/(d-2)} \in \left(\frac{1-\sin\phi}{\cos\phi}, \frac{1}{\cos\phi}\right), \text{ and}$$
$$\cos\phi > \left\{-(d-3)[(-z)^{1/(d-2)} + (-z)^{-1/(d-2)}] + \sqrt{\Delta}\right\}/4,$$
where $\Delta = (d-3)^2[(-z)^{1/(d-2)} + (-z)^{-1/(d-2)}]^2 + 16(d-4)$.

- When $A < 0$, we already knew that $-\frac{x_2}{x_3} > 1$. When $z > 0$ and d is odd, we have $x_2 < 0 < x_1 < x_3 = \cdots = x_d$. We can show that in this case $x_1 + x_2 = [(d-2)A + 2n^2]/(1+\lambda) < 0$. Thus, $-\frac{x_2}{x_1} > 1$.

- When $0 < n^2 < A$, we have $x_3 = \cdots = x_d < 0 < x_2 < x_1$. This is not an important case. But we can still find that $-\frac{x_3}{x_2} > 1$.

8 Week of October 27 - November 2

Define

$$\nu = -\frac{\max\{\text{negative eigenvalues}\}}{\min\{\text{positive eigenvalues}\}},$$

then $\nu < 1$ implies that the leading stable direction is stronger than the leading unstable direction; therefore, it is the bad situation we want to study.

In the G1T2 case when the critical points are in the form of $\alpha_1 \neq \alpha_2 = \cdots = \alpha_d$, I checked last week's calculations and looked closer into the results, and found that the entire parameter space can be divided to:

Case 1: $(-z)^{1/(d-2)} \in (-1, 0) \cup (\cos\phi, 1)$, we have $A < 0$ and $n^2 - A > 0$. Then $x_1 > 0$, $x_3 = \cdots = x_d > 0$, and $x_2 < 0$. We can show that $-\frac{x_2}{x_3} > 1$.

- If $x_2 < 0 < x_1 < x_3$, then $\nu = -\frac{x_2}{x_1} > -\frac{x_2}{x_3} > 1$.
- If $x_2 < 0 < x_3 < x_1$, then $\nu = -\frac{x_2}{x_3} > 1$.

Case 2: $(-z)^{1/(d-2)} \in (0, \frac{1-\sin\phi}{\cos\phi})$, then we have $0 < n^2 < A$ and $n^2 - A < 0$. In this case, $x_3 = \cdots = x_d < 0 < x_2 < x_1$. We find that $\nu = -\frac{x_3}{x_2} > 1$.

Case 3: $(-z)^{1/(d-2)} \in (\frac{1-\sin\phi}{\cos\phi}, \cos\phi)$, we have $0 < A < n^2$ and $n^2 - A > 0$. We know that $x_2 < 0$, $x_1 > 0$ and $x_3 = \cdots = x_d > 0$. We can show that $-\frac{x_2}{x_1} < 1$.

- When $d = 3, 4$ or 5, it is easy to see that $-\frac{x_2}{x_3} < 1$. Thus $\nu < 1$.
- When $d \geq 6$, in order to have $-\frac{x_2}{x_3} < 1$, we need the restriction that

$$\cos\phi \in \left(\frac{(d-3)[(-z)^{-1/(d-2)} + (-z)^{1/(d-2)}] - \sqrt{\Delta}}{4}, 1\right],$$

where $\Delta = (d-3)^2[(-z)^{-1/(d-2)} + (-z)^{1/(d-2)}]^2 - 16(d-4)$.
$(-z)^{1/(d-2)} \in (\frac{1-\sin\phi}{\cos\phi}, \cos\phi)$ is equivalent to $\cos\phi \in \left((-z)^{1/(d-2)}, \frac{2(-z)^{1/(d-2)}}{1+(-z)^{2/(d-2)}}\right)$.
We can find the parameter region where $\nu < 1$ is

$$\left((-z)^{1/(d-2)}, \frac{2(-z)^{1/(d-2)}}{1+(-z)^{2/(d-2)}}\right) \cap \left(\frac{(d-3)[(-z)^{-1/(d-2)} + (-z)^{1/(d-2)}] - \sqrt{\Delta}}{4}, 1\right]$$
$$= \left(\frac{(d-3)[(-z)^{-1/(d-2)} + (-z)^{1/(d-2)}] - \sqrt{\Delta}}{4}, \frac{2(-z)^{1/(d-2)}}{1+(-z)^{2/(d-2)}}\right).$$

From the figures on the next page, we can see that ν gets smaller when the parameters are closer from below to the curve $\cos\phi = 2(-z)^{1/(d-2)}/(1 + (-z)^{2/(d-2)})$ (equivalent to $(-z)^{1/(d-2)} = (1 - \sin\phi)/\cos\phi$).

But it seems that when d gets larger, the parameter region for $\nu < 1$ gets smaller.

When $d = 3$ or 4, the parameter region of case 1 shows $\nu < 1$ as well.

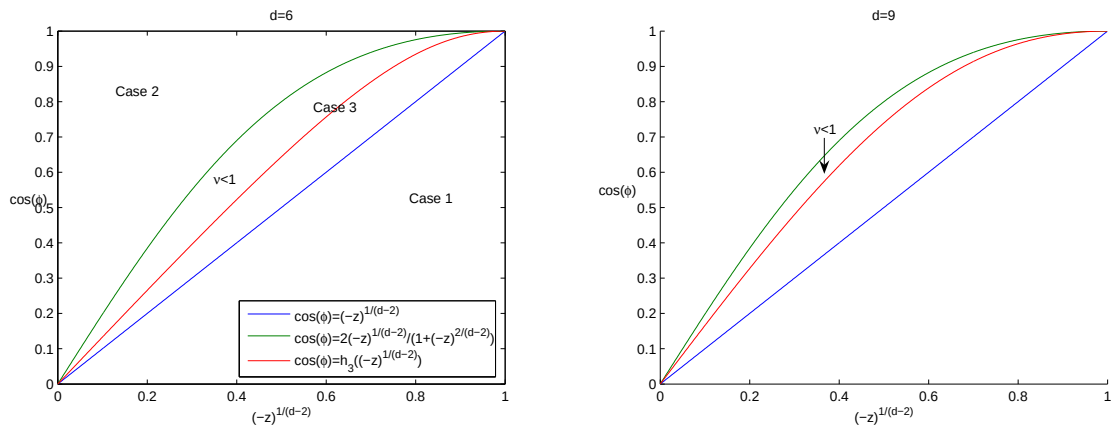


Figure 1: The parameter space with $z < 0$ and $\phi \in (0, \pi/2)$.

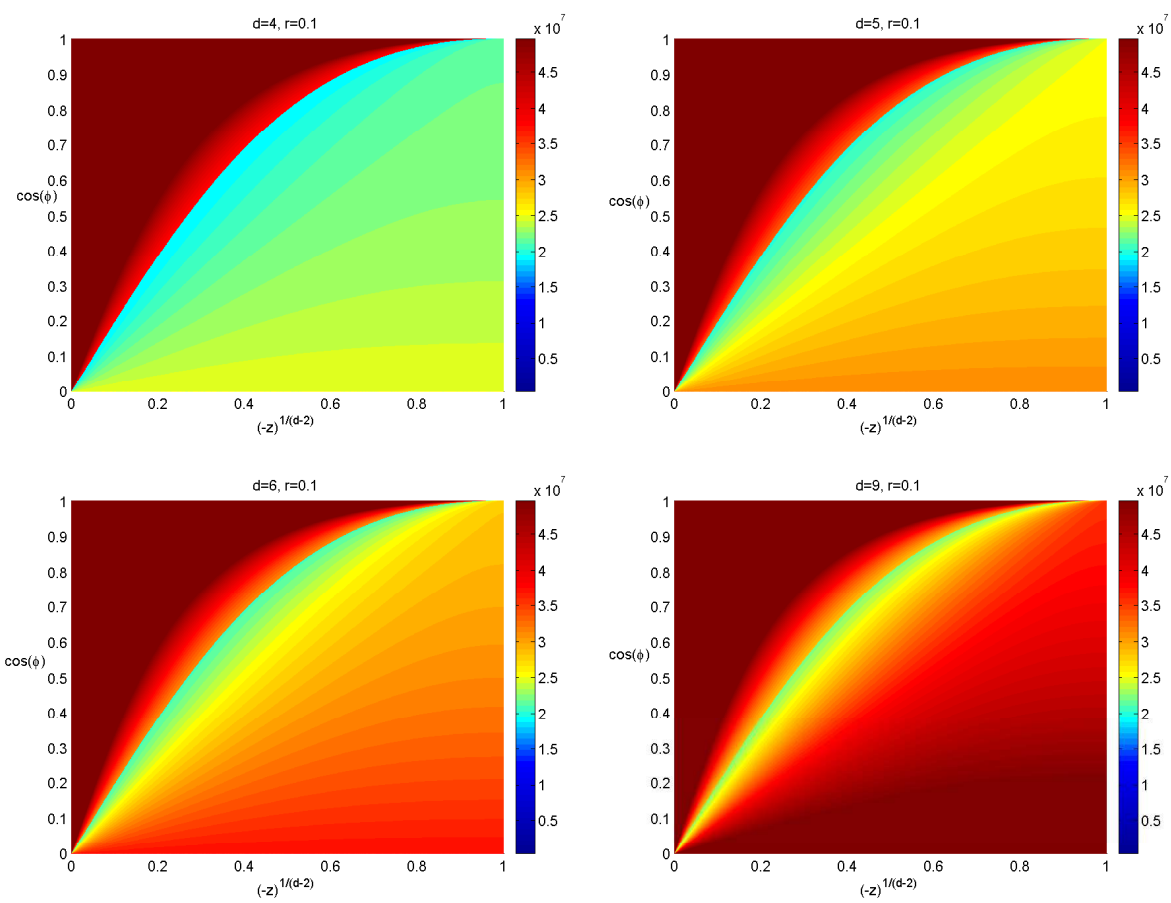


Figure 2: The plots of ν under different parameter values. The color bar is wrong. $\nu = 1$ corresponds to the yellow color. The light blue area corresponds to $\nu \approx 0.5$.

9 Week of November 3 - November 9

This week I worked on numerically finding the parameter regions where $\nu < 1$ at the saddles with the form $\alpha_1 = \cdots = \alpha_m \neq \alpha_{m+1} = \cdots = \alpha_d$ in the G1T2 case.

1. When d is odd and $m = 2$, I observe no regions with $\nu < 1$. The plots look the same

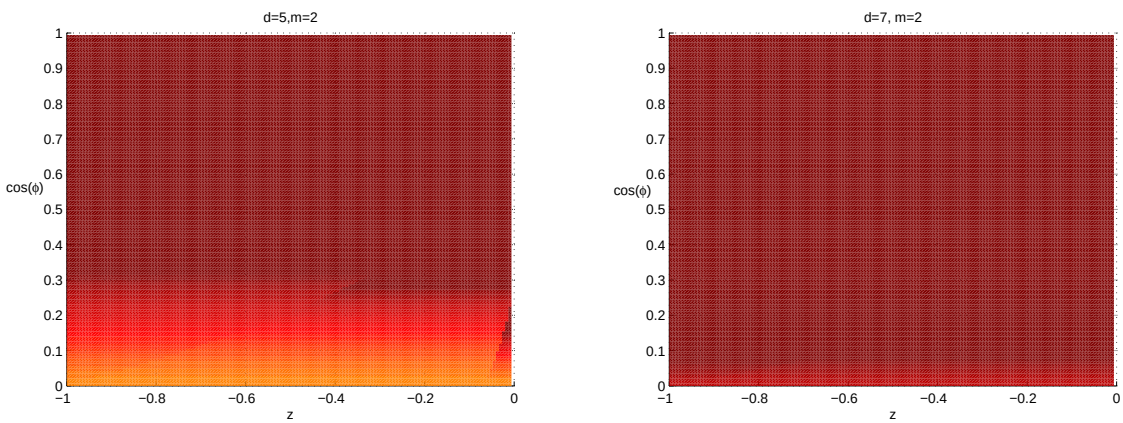


Figure 3: The parameter space with $z < 0$ and $\phi \in (0, \pi/2)$.

   for $z > 0$.

2. When d is odd and $m = 3$, there exists some small parameter regions with $z > 0$ where ν is very close to 0.

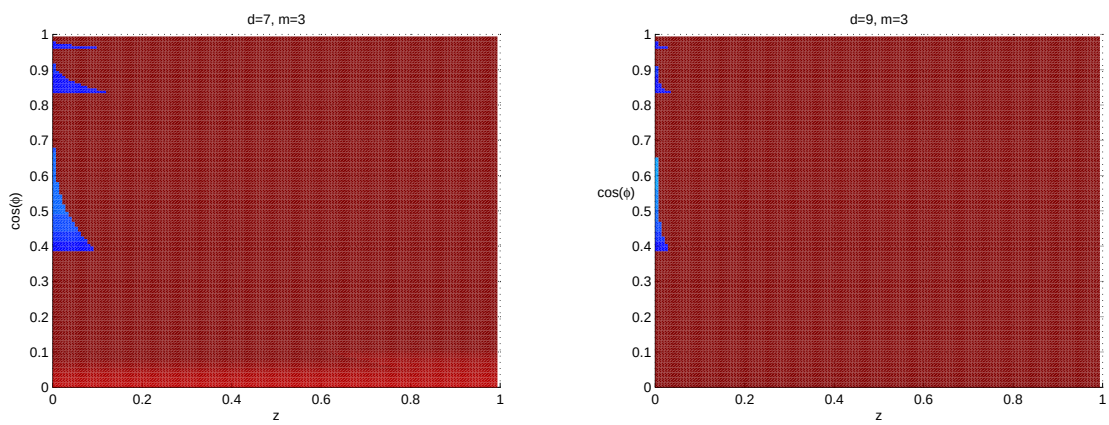
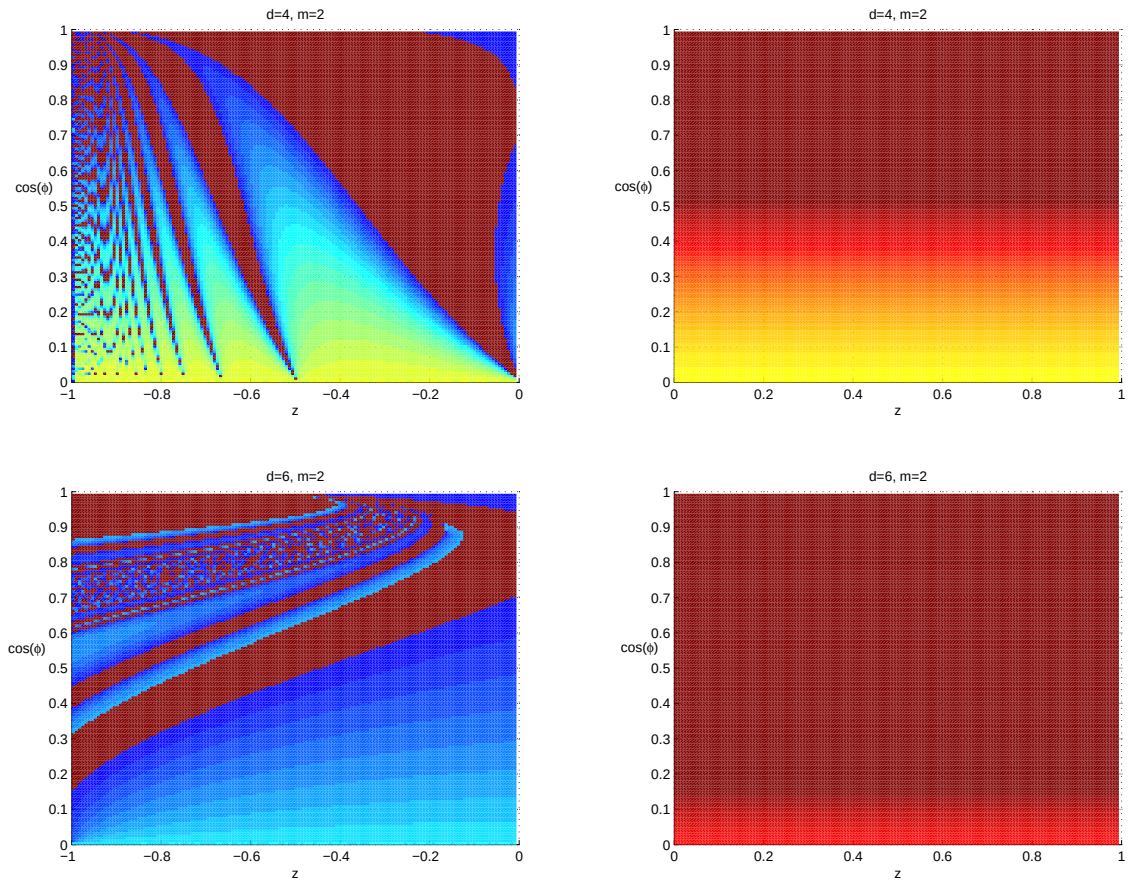


Figure 4: The parameter space with $z > 0$ and $\phi \in (0, \pi/2)$.

3. When d is even and $m = 2$:

   - For $z < 0$, the saddle points do not exist for all ϕ and z. But we do observe some regions with $\nu < 1$.

- For $z > 0$, we observe $\nu = 1$ for $d = 4$ but not $\nu < 1$. For larger d, we have $\nu > 1$.



10 Week of October 10 - November 16

Using

$$\nu = -\frac{\min\{\text{negative eigenvalues}\}}{\min\{\text{positive eigenvalues}\}},$$

I ran the computations again and obtained the following plots.

1. For d is odd, any m and any z, the critical point always exists (Fig. 5).
2. When d is even and m is odd, the critical point exists when $z < 0$ (Fig. 6).
3. When d is even and m is even, the critical point exists when $z > 0$ (Fig. 7) and when $z < 0$ with certain restrictions (Fig. 8).

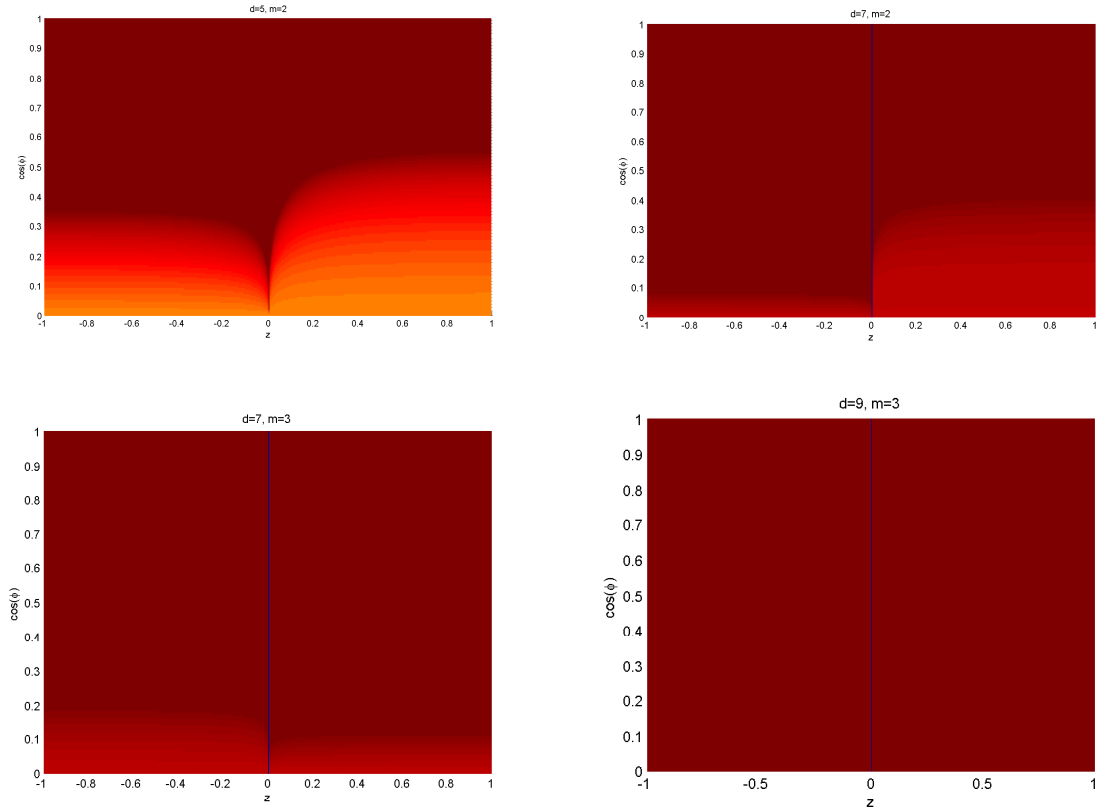


Figure 5: Plot of ν in the parameter space when $d = 5, 7$ or 9 and $m = 2$ or 3. The vertical line at $z = 0$ corresponds to the case where at lease one of the eigenvalues at the saddle point is 0. There are no regions where $\nu < 1$.

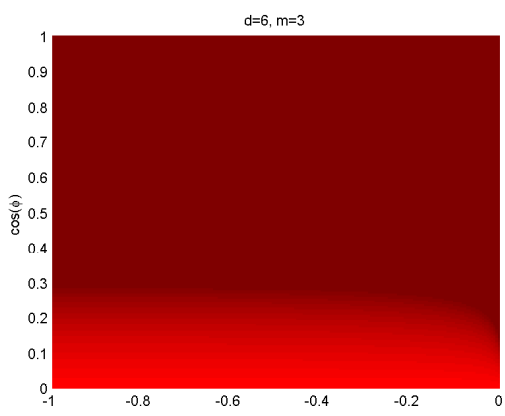


Figure 6: When $d = 6$, $m = 3$ and $z < 0$, there are no regions for $\nu < 1$. For larger d, the plot looks similar.

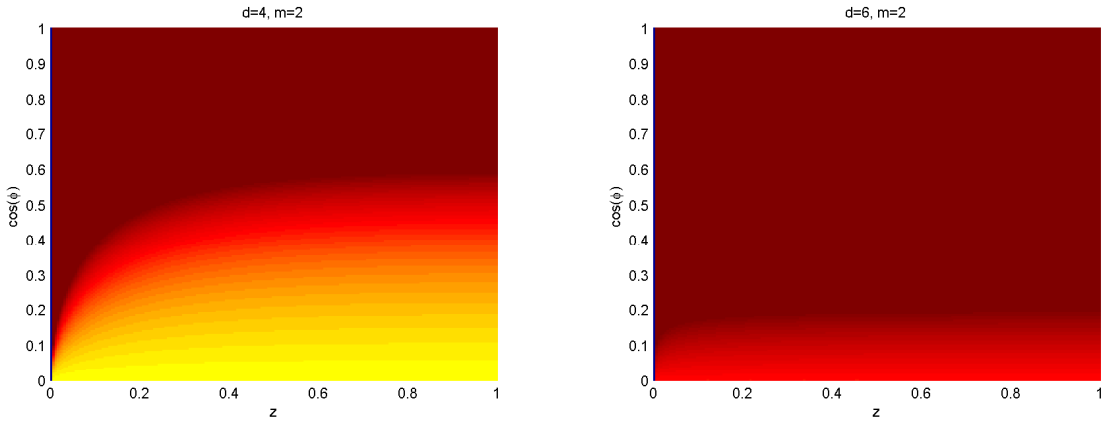


Figure 7: For $d = 4$ or 6, $m = 2$ and $z > 0$, we see no region for $\nu < 1$. The yellow color corresponds to $\nu = 1$ on the left picture.

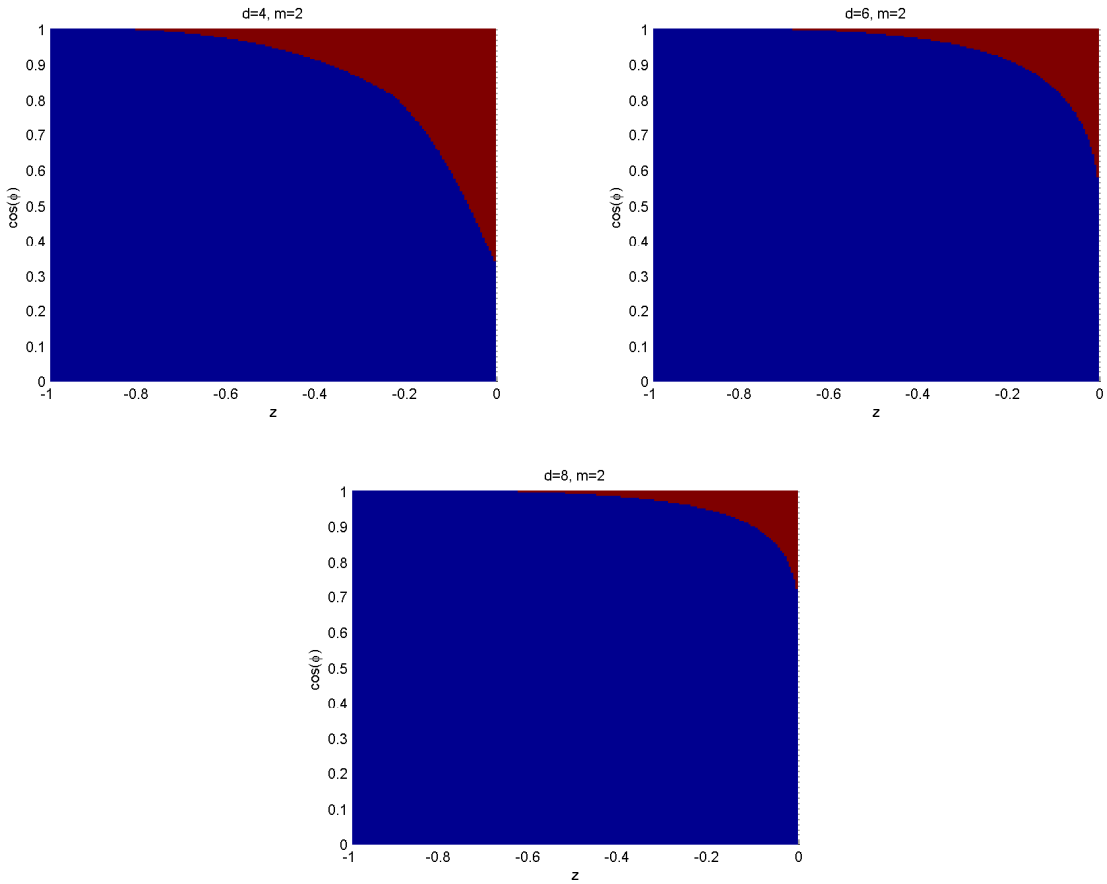


Figure 8: For $d = 4$ or 6, $m = 2$ and $z < 0$, blue regions are where the critical point does not exist. For where the critical points exist, we have $\nu > 1$.