Journal of Xue Gong for Spring semester 2016

May 16, 2016

1 January 11 - 19

This week I worked on the G2T2 case, and looked at if it is possible for $\alpha_1 \neq \alpha_2 = \cdots = \alpha_d$ and $\alpha_2 \neq \alpha_1 = \alpha_3 = \cdots = \alpha_d$ to happen.

2 January 20 - 26

I kept working on the G2T2 case. I tried to find more solutions to the equilibrium points to the gradient flow, but my approach was wrong.

3 January 27 - February 2

I continued to study the G2T2 case. And I started to edit the tensor chapter in my dissertation.

4 February 3 - 9

I edited the tensor chapter, and studied the saddle value ν when $\alpha_1 = \cdots = \alpha_m \neq \alpha_{m+1} = \cdots = \alpha_d$ for $2 \leq m \leq d-2$.

5 February 10 - 16

I proved that $\nu > 1$ when $2 \leq m \leq d-2$, and related this situation to $m = 1$. The same results for $m = 1$ can be found when using the formulas in the case of $2 \leq m \leq d-2$ and setting $m = 1$. So we can combine the two cases together.

6 February 17 - 23

I wrote the results from last week into my dissertation and made more corrections on typos and grammar mistakes.

7 February 24 - March 1

I submitted my dissertation to the committee members, and started to edit the dissertation according to the feedback from the committee members.

8 March 2 - 8

I continued to work on editing the dissertation, and also prepared for the dissertation defense.

9 March 9 - 15

I passed the dissertation defense on March 10. Then I did more editing of the dissertation so that it can be finally submitted. I also prepared a poster for the Pi Day poster presentations on my results in the G1T2 case of tensor approximations.

10 March 16 - 22

I started to look at a special G1T2 case when $\phi = \pi/2$ and $\alpha_1 \neq \alpha_2 = \cdots = \alpha_d$, there is a non-hyperbolic equilibrium point of the gradient flow. We want to show that this is still a saddle point.

11 March 23 - 29

In the G1T2 case, we know that when $n(\vec{\alpha}) = 0$, the regularized error function reaches its maximum, $E_\lambda(G_1) = 1$. In this situation,

$$\prod_{i=1}^{d} \cos(\alpha_i) + z \prod_{i=1}^{d} \cos(\alpha_i - \phi) = 0.$$

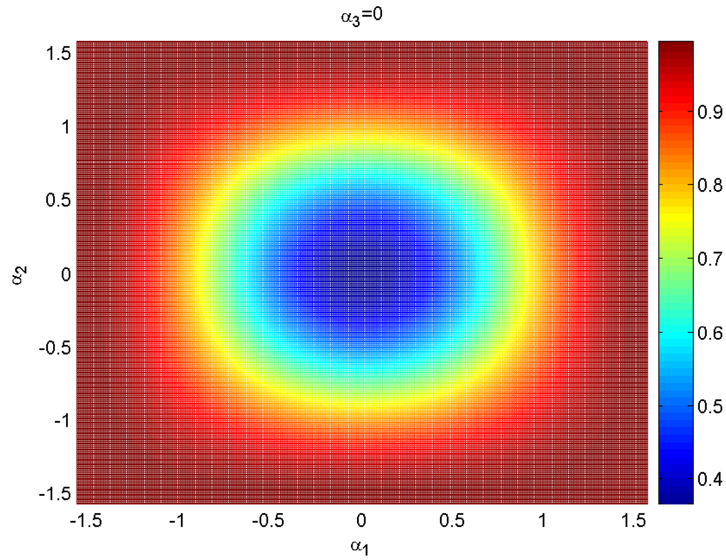
The solutions are

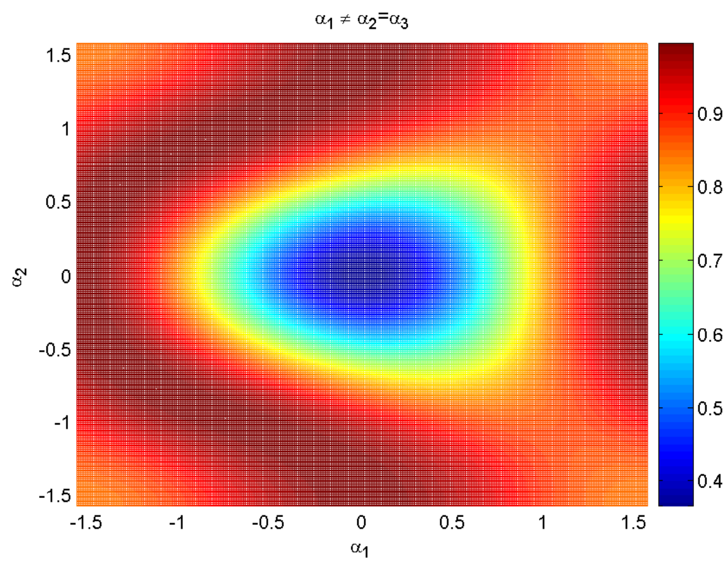
1. $\alpha_i = \pi/2$ and $\alpha_j = \phi - \pi/2$ for some $i \neq j$.
2. If $\alpha_i \neq \pi/2$ for $1 \leq i \leq d-1$ and $\alpha_d \neq \phi - \pi/2$, then

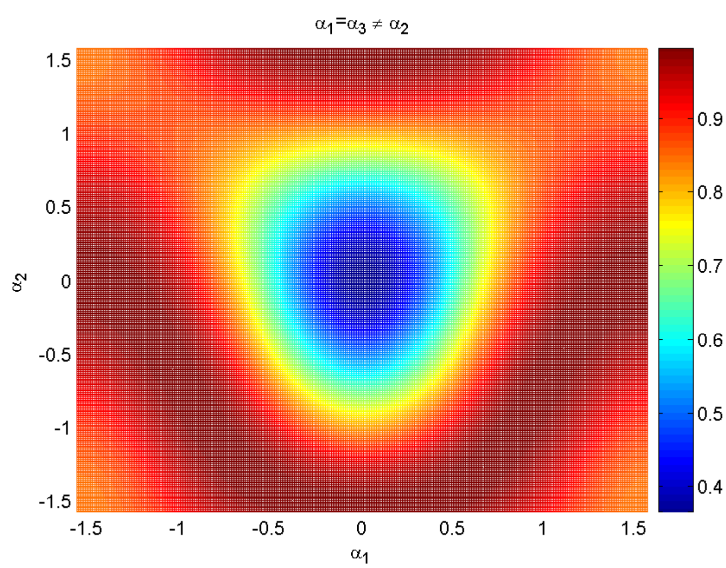
$$\frac{\cos(\alpha_d)}{\cos(\alpha_d - \phi)} = -z \prod_{i=1}^{d-1} \frac{\cos(\alpha_i - \phi)}{\cos(\alpha_i)}.$$

   Therefore, there is a curve consisting of maximum points where α_d is a function of $\alpha_1, \ldots, \alpha_{d-1}$, given that $\alpha_i \notin \{\pi/2, \phi - \pi/2\}$ for all i and $z \neq 0$.

   In particular, when $\phi = \pi/2$, we found that this curve satisfies $\cot(\alpha_d) = -z \prod_{i=1}^{d-1} \tan(\alpha_i)$. For $d = 3$, $\phi = \pi/2$ and $z = 0.5$, we plot the error function:


α3=0
1.5
1
0.5
0
-0.5
-1
-1.5
α2
-1.5 -1 -0.5 0 0.5 1 1.5
α1
0.9
0.8
0.7
0.6
0.5
0.4


α1 ≠ α2=α3
1.5
1
0.5
0
-0.5
-1
-1.5
α2
-1.5 -1 -0.5 0 0.5 1 1.5
α1
0.9
0.8
0.7
0.6
0.5
0.4


α1=α3 ≠ α2
1.5
1
0.5
0
-0.5
-1
-1.5
α2
-1.5 -1 -0.5 0 0.5 1 1.5
α1
0.9
0.8
0.7
0.6
0.5
0.4

12 March 30 - April 5

We consider the maximum points of the regularized error function as last week. In case 2 where $\alpha_i \notin \{\pi/2, \phi - \pi/2\}$, we have

$$\frac{\cos(\alpha_d)}{\cos(\alpha_d - \phi)} = -z \prod_{i=1}^{d-1} \frac{\cos(\alpha_i - \phi)}{\cos(\alpha_i)}. \tag{1}$$

This can be simplified as (when $\alpha_i \neq \pi/2$ for all i)

$$(\cos\phi)^d + \sum_{i=1}^{d} \tan\alpha_i (\cos\phi)^{d-1} \sin\phi + \sum_{i<j} \tan\alpha_i \tan\alpha_j (\cos\phi)^{d-2} (\sin\phi)^2$$
$$+ \sum_{i<j<k} \tan\alpha_i \tan\alpha_j \tan\alpha_k (\cos\phi)^{d-3} (\sin\phi)^3 + \cdots + \prod_{i=1}^{d} \tan\alpha_i (\sin\phi)^d = -\frac{1}{z}.$$

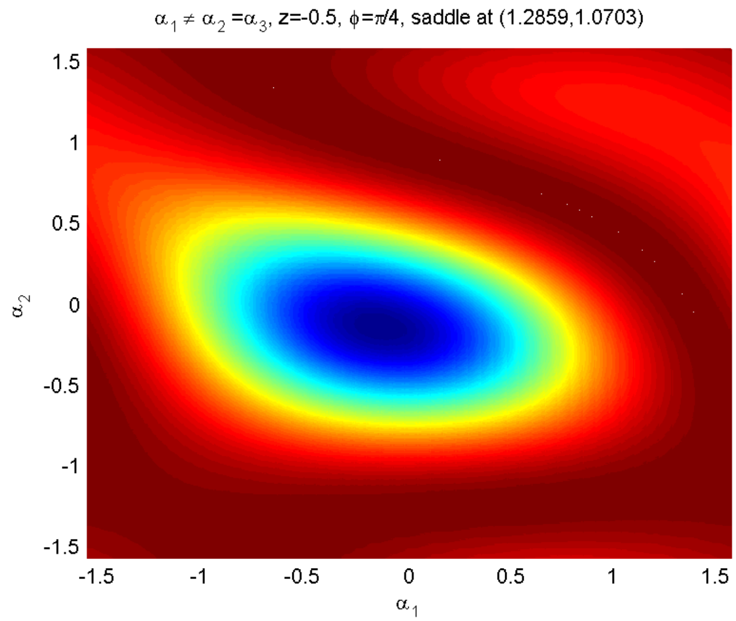
We can find that the left-hand side of (1) is one-to-one in $(-\pi/2, \pi/2)$ because

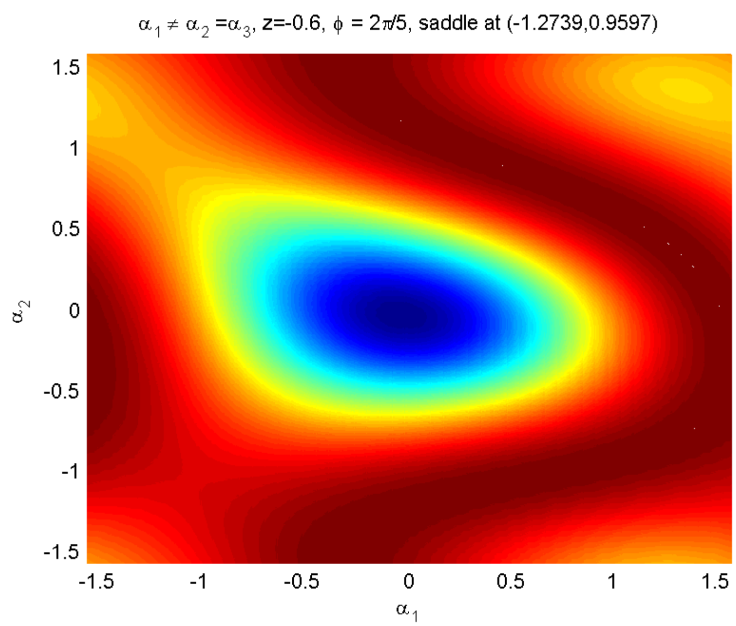
$$\frac{\cos(\alpha_d)}{\cos(\alpha_d - \phi)} = \frac{1}{\cos\phi + \tan\alpha_d \sin\phi}.$$

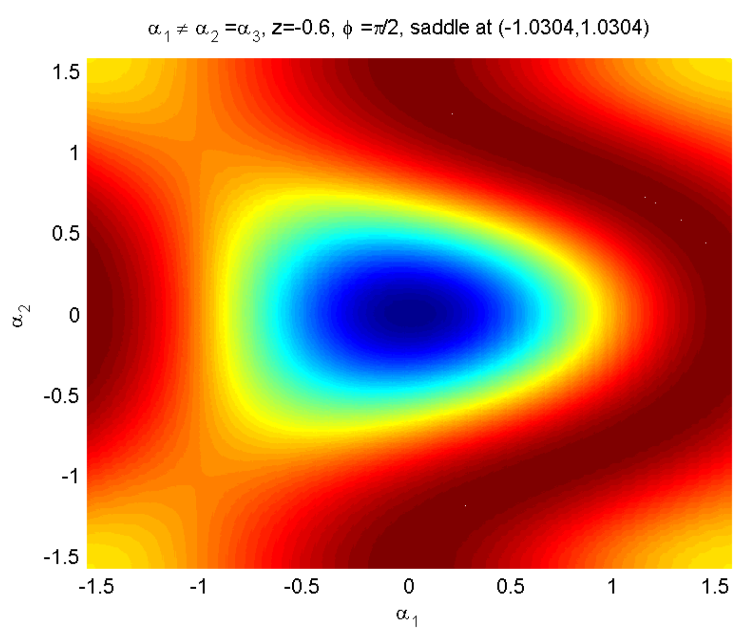
Therefore, there is only one sheet of maximum points.

When $d = 3$, at the saddle point, there is a one-dimensional unstable manifold and a two-dimensional stable manifold. One unstable direction goes to the minimum point in the symmetric set; while the other opposite direction goes to another local minimum point on the symmetric set (the diagonal of the following plots).

When $\phi = \pi/2$, there is only one saddle point in the form of $\alpha_1 \neq \alpha_2 = \alpha_3$ (three saddles considering the permutations). We see the same unstable direction even though this is a non-hyperbolic saddle.


α1 ≠ α2 =α3, z=-0.5, ϕ=π/4, saddle at (1.2859,1.0703)
α2
α1
1.5
1
0.5
0
-0.5
-1
-1.5


α1 ≠ α2 =α3, z=-0.6, ϕ = 2π/5, saddle at (-1.2739,0.9597)
α2
α1
1.5
1
0.5
0
-0.5
-1
-1.5


α1 ≠ α2 =α3, z=-0.6, ϕ =π/2, saddle at (-1.0304,1.0304)
α2
α1
1.5
1
0.5
0
-0.5
-1
-1.5

13 April 6 - 12

This week, I started to put together a paper *The Optimization Landscape for Fitting a Rank-2 Tensor with a Rank-1 Tensor* with Dr. Mohlenkamp.

Because the saddle value $\nu \in (0, \infty)$, we map it to $\tilde{\nu} \in (0, 1)$ by defining

$$\tilde{\nu} = \frac{\nu}{1+\nu}.$$

In this way, $\tilde{\nu} < \frac{1}{2}$ indicates that the strongest unstable direction is weaker than the weakest stable direction. We obtain the following figures.

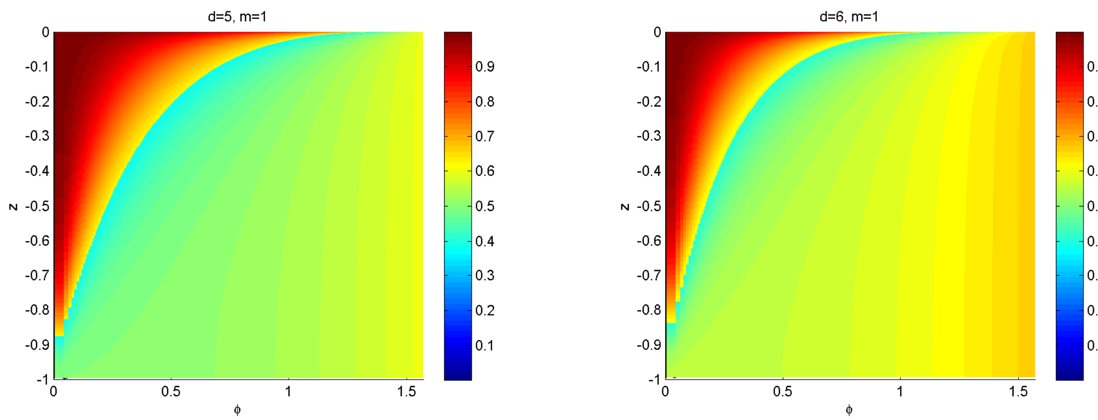


Figure 1: Plots of $\tilde{\nu}$ when $d = 5$ and $d = 6$. The saddle points satisfy $\alpha_1 \neq \alpha_2 = \cdots = \alpha_d$.

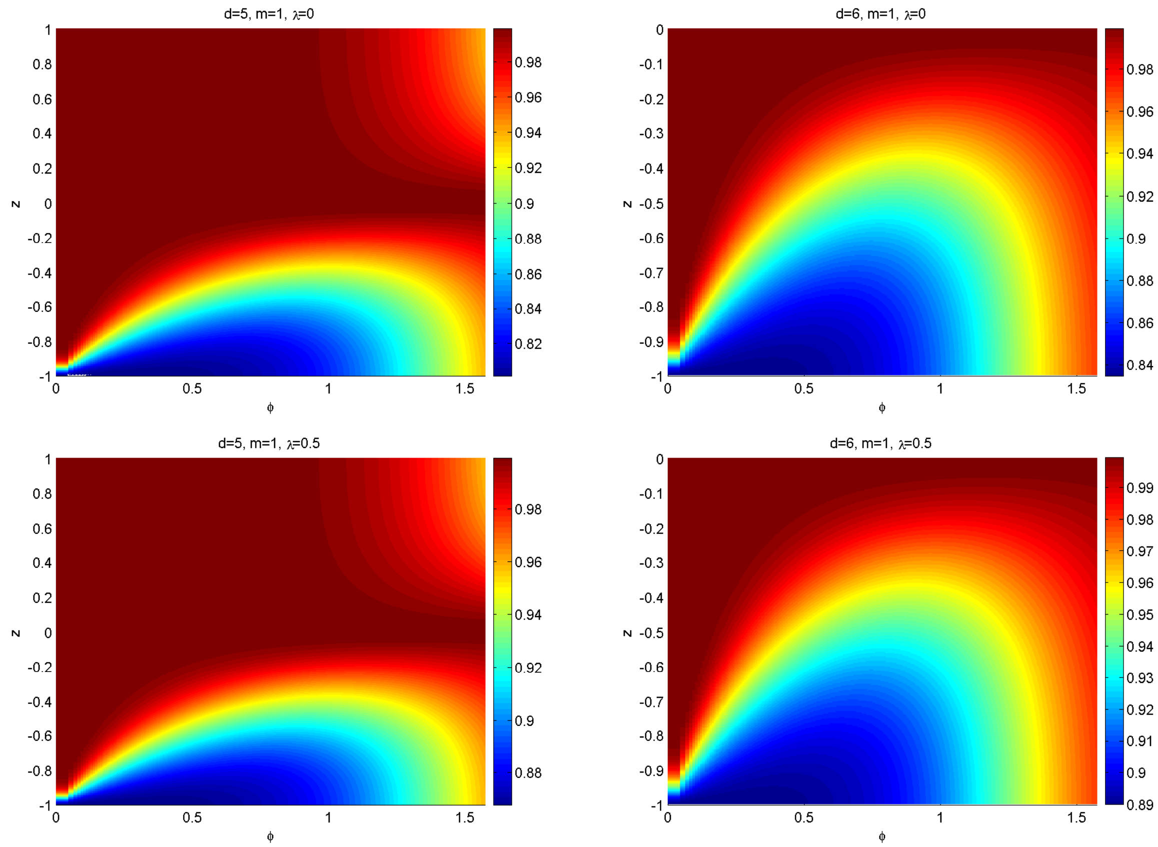


Figure 2: Plots of $E_0(G_1)$ at the saddle points when $d = 5$ and $d = 6$. The saddle points satisfy $\alpha_1 \neq \alpha_2 = \cdots = \alpha_d$.

It looks like the errors at the saddle points are all relatively large. There are some points where it looks like $E_\lambda(G_1) = 1$. These are not the maximum points, so it is possible that the error is very close to 1.

14 April 13 - 19

This week I am working on the *The Optimization Landscape for Fitting a Rank-2 Tensor with a Rank-1 Tensor* paper. I added a subsection about the maximum points, and try to combine the $m = 1$ case and the $m \geq 2$ case together.