Department of Mathematical Sciences

TMA4500 - Industrial Mathematics, Specialization Project

Optimization using second order information on the Symplectic Stiefel manifold

Author:
Ole Gunnar Røsholt Hovland

Date

Table of Contents

List of Figures

List of Tables

1 Introduction

Hi! Welcome to this LaTeX-template. I will here aim to introduce you to, as well as motivate you to learn more about the features available in LaTeX through Overleaf. Many of the features you will come across in this template are not necessarily relevant to you at this point in time, and some will most likely seem way too advanced. However, keep in mind that you are not expected to understand everything at once either.

I hope that you, with the assistance of what I provide you with here, are able to make your own LaTeX-templates containing your personal preferences. You may do it by directly changing variables in this template, or you may create a brand new containing only carefully selected features of your own.

As a final note, I want to wish you the best of luck learning LaTeX, but do keep in mind that this template is only scratching the surface.

2 Example Section

In physics, the Navier–Stokes equations, named after Claude-Louis Navier and George Gabriel Stokes, describe the motion of viscous fluid substances.

(2.1) shows the incompressible Navier-Stokes equations using tensor notation.

$$\frac{\partial u_i}{\partial t} + u_j \frac{\partial u_i}{\partial x_j} = -\frac{1}{\rho}\frac{\partial P}{\partial x_i} + \nu \frac{\partial^2 u_i}{\partial x_j^2} \tag{2.1}$$

Figure 1: Caption written below figure.

Source: Insert image source here

3 Theory

3.1 Basic definitions

Definition 1 (Smooth manifold). *Given a topological space $\mathcal{M}$, it is a topological manifold of dimension n if*

1. *$\mathcal{M}$ is a Hausdorff space: for every pair of $p, q \in \mathcal{M}$, we can always find to disjoint subsets of $\mathcal{M}$, U and V, such that $p \in U$ and $q \in V$.*
2. *$\mathcal{M}$ is second-countable: the topology of $\mathcal{M}$ has a countable basis.*

3. $\mathcal{M}$ *is locally Euclidean of dimension* n*: for each* p *we have an open subset* $U \subseteq \mathcal{M}$ *containing* p*, and an open subset* $\hat{U} \in \mathbb{R}^n$ *such that there exists a homeomorphism* $\phi : U \to \hat{U}$*.*

If we in addition have a notion of smoothness, meaning that the notion of differentiability is well-defined, we call $\mathcal{M}$ *a smooth manifold.*

For the rest of this paper we denote $\mathcal{M}$ as being a Riemannian manifold.

3.2 Right-invariant framework

(many of the usual SpSt metrics do not have geodesics, therefore...)(We need a proper metric for $\mathrm{SpSt}(2n, 2k)$ in the sense that we need a metric that allows us to perform optimization on this manifold. To this end we need a metric that makes it possible to derive geodesics, as opposed to (BZ sources). The following metric defined in BZ fulfils our criteria.)

The goal for this section is to use a right-invariant metric defined on the symplectic group to define an appropriate metric on the symplectic Stiefel manifold. Define the point-wise right-invariant metric on $\mathrm{Sp}(2n, \mathbb{R})$ as the mapping $g_p : T_p\mathrm{Sp}(2n, \mathbb{R}) \times T_p\mathrm{Sp}(2n, \mathbb{R}) \to \mathbb{R}$,

$$g_p(X_1, X_2) := \frac{1}{2}\,\mathrm{tr}((X_1 M^+)^T X_2 M^+), \quad X_1, X_2 \in T_p\mathrm{Sp}(2n, \mathbb{R}). \tag{3.1}$$

It is right-invariant in the sense that $g_{pq}(X_1 q, X_2 q) = \frac{1}{2}\mathrm{tr}((X_1 q q^+ p^+)^T X_2 q q^+ p^+) = g_p(X_1, X_2)$ for all $p \in \mathrm{Sp}(2n, \mathbb{R})$.

We will now use a *horizontal lift* to define a metric on $\mathrm{SpSt}(2n, 2k)$ through 3.1. Split $T_p\mathrm{Sp}(2n, \mathbb{R})$ into to parts: the horizontal- and vertical part, with respect to g_p^{Sp} and π:

$$T_p\mathrm{Sp}(2n, \mathbb{R}) = \mathrm{Ver}_p^\pi \oplus \mathrm{Hor}_p^\pi\, \mathrm{Sp}(2n, \mathbb{R}). \tag{3.2}$$

The point-wise right-invariant Riemannian metric on $\mathrm{SpSt}(2n, 2k)$ is defined as the mapping $g_p : T_p\mathrm{SpSt}(2n, 2k) \times T_p\mathrm{SpSt}(2n, 2k) \to \mathbb{R}$, $g_p(X_1, X_2) := g_p^{\mathrm{Sp}}((X_1)_p^{\mathrm{hor}}, (X_2)_p^{\mathrm{hor}})$. More explicitly

$$g_p(X_1, X_2) = \mathrm{tr}\left(X_1^T \left(I_{2n} - \frac{1}{2} J_{2n}^T p (p^T p)^{-1} p^T J_{2n} \right) X_2 (p^T p)^{-1} \right), \tag{3.3}$$

for $X_1, X_2 \in T_p\mathrm{SpSt}(2n, 2k)$. For this metric, π denotes a Riemannian submersion.

Now that we have chosen a metric, we can justify a choice for a Riemannian gradient.

Proposition 1. *Given a function* $f : \mathrm{SpSt}(2n, 2k) \to \mathbb{R}$*, the Riemannian gradient with respect to* g_p *is given by*

$$\mathrm{grad}\, f(p) = \nabla f(p) p^T p + J_{2n} p (\nabla f(p))^T J_{2n} p, \tag{3.4}$$

where $\nabla f(p)$ *is the Euclidean gradient of a smooth extension around* $p \in \mathrm{SpSt}(2n, 2k)$ *in* $\mathbb{R}^{2n \times 2k}$ *at* p*.*

Proof. We can see that this is the Riemannian gradient by the following two observations stated in **BZ**, which we verify here.

Firstly, gradient must be in $T_p\mathrm{SpSt}(2n, 2k)$, which means by ref ?? that $0 = p^+ \,\mathrm{grad}\, f(p) + (\mathrm{grad}\, f(p))^+ p$ so

$$p^\mathrm{T} J \nabla f(p) p^\mathrm{T} p + p^\mathrm{T} J J p (\nabla f(p))^\mathrm{T} J p + p^\mathrm{T} p (\nabla f(p))^\mathrm{T} J p + p^\mathrm{T} J^\mathrm{T} \nabla f(p) p^\mathrm{T} J^\mathrm{T} J p \tag{3.5}$$

$$= -p^\mathrm{T} J^\mathrm{T} \nabla f(p) p^\mathrm{T} J^\mathrm{T} J p - p^\mathrm{T} p (\nabla f(p))^\mathrm{T} J p + p^\mathrm{T} p (\nabla f(p))^\mathrm{T} J p + p^\mathrm{T} J^\mathrm{T} \nabla f(p) p^\mathrm{T} J^\mathrm{T} J p = 0 \tag{3.6}$$

! Or should I mark where I did the simplification for clarity:

$$
\begin{aligned}
&p^{\mathrm{T}} J \nabla f(p) p^{\mathrm{T}} p + p^{\mathrm{T}} \cancelto{-I_{2n}}{JJ} p (\nabla f(p))^{\mathrm{T}} J p + p^{\mathrm{T}} p (\nabla f(p))^{\mathrm{T}} J p + p^{\mathrm{T}} J^{\mathrm{T}} \nabla f(p) p^{\mathrm{T}} \cancelto{I_{2n}}{J^{\mathrm{T}} J} p && (3.7)\\
&= -p^{\mathrm{T}} J^{\mathrm{T}} \nabla f(p) p^{\mathrm{T}} J^{\mathrm{T}} J p - p^{\mathrm{T}} p (\nabla f(p))^{\mathrm{T}} J p + p^{\mathrm{T}} p (\nabla f(p))^{\mathrm{T}} J p + p^{\mathrm{T}} J^{\mathrm{T}} \nabla f(p) p^{\mathrm{T}} J^{\mathrm{T}} J p = 0 && (3.8)
\end{aligned}
$$

where we have used $JJ = -J^{\mathrm{T}} J = -I_{2n}$ and $J^{\mathrm{T}} = -J$.

Secondly, the gradient also has to satisfy $g_p(\operatorname{grad} f(p), X) = \mathrm{d} f_p(X) = \operatorname{tr}((\nabla f(p))^{\mathrm{T}} X)$ for all $X \in T_p\mathrm{SpSt}(2n, 2k)$:

$$
g_p(\operatorname{grad} f(p), X) = \operatorname{tr}\big((p^{\mathrm{T}} p (\nabla f(p))^{\mathrm{T}} + p^{\mathrm{T}} J^{\mathrm{T}} \nabla f(p) p^{\mathrm{T}} J^{\mathrm{T}})(I_{2n} - \tfrac{1}{2} G) X (p^{\mathrm{T}} p)^{-1}\big),
$$

where $G := J^{\mathrm{T}} p (p^{\mathrm{T}} p)^{-1} p^{\mathrm{T}} J$. Expanding this expression we obtain

$$
\begin{aligned}
&\operatorname{tr}\big(\cancel{p^{\mathrm{T}} p} (\nabla f(p))^{\mathrm{T}} X \cancel{(p^{\mathrm{T}} p)^{-1}}\big) - \tfrac{1}{2} \operatorname{tr}\big(\cancel{p^{\mathrm{T}} p} (\nabla f(p))^{\mathrm{T}} G X \cancel{(p^{\mathrm{T}} p)^{-1}}\big) && (3.9)\\
&+ \operatorname{tr}\big(p^{\mathrm{T}} J^{\mathrm{T}} \nabla f(p) p^{\mathrm{T}} J^{\mathrm{T}} X (p^{\mathrm{T}} p)^{-1}\big) - \tfrac{1}{2} \operatorname{tr}\big(p^{\mathrm{T}} J^{\mathrm{T}} \nabla f(p) p^{\mathrm{T}} J^{\mathrm{T}} G X (p^{\mathrm{T}} p)^{-1}\big), && (3.10)
\end{aligned}
$$

where the cancellations used the fact that the trace is invariant under circular shifts. Noting that the first term is by definition $\mathrm{d} f_p(X)$, and inserting the definition of G, the expression becomes

$$
\begin{aligned}
&\mathrm{d} f_p(X) - \tfrac{1}{2} \operatorname{tr}\big((\nabla f(p))^{\mathrm{T}} J^{\mathrm{T}} p (p^{\mathrm{T}} p)^{-1} p^{\mathrm{T}} J X\big) + \operatorname{tr}\big(p^{\mathrm{T}} J^{\mathrm{T}} \nabla f(p) p^{\mathrm{T}} J^{\mathrm{T}} X (p^{\mathrm{T}} p)^{-1}\big) && (3.11)\\
&- \tfrac{1}{2} \operatorname{tr}\big(p^{\mathrm{T}} J^{\mathrm{T}} \nabla f(p) p^{\mathrm{T}} \cancelto{-I_{2n}}{J^{\mathrm{T}} J^{\mathrm{T}}} p (p^{\mathrm{T}} p)^{-1} p^{\mathrm{T}} \cancelto{-J^{\mathrm{T}}}{J} X (p^{\mathrm{T}} p)^{-1}\big). && (3.12)
\end{aligned}
$$

We notice that after cancelling $p^{\mathrm{T}} p (p^{\mathrm{T}}) p^{-1}$ in the last term, the trace is equal to the second to last term. Now focusing on the second term: utilizing both the fact that for any matrix, A, (a) $\operatorname{tr}(A) = \operatorname{tr}(A^{\mathrm{T}})$, (b) the cyclic property of the trace, and (c) $J = -J^{\mathrm{T}}$, we get that

$$
\begin{aligned}
\tfrac{1}{2} \operatorname{tr}\big((\nabla f(p))^{\mathrm{T}} J^{\mathrm{T}} p (p^{\mathrm{T}} p)^{-1} p^{\mathrm{T}} J X\big) &\overset{(1)}{=} \tfrac{1}{2} \operatorname{tr}\big(X^{\mathrm{T}} J^{\mathrm{T}} p (p^{\mathrm{T}} p)^{-1} p^{\mathrm{T}} J \nabla f(p)\big) && (3.13)\\
&\overset{(2),(3)}{=} -\tfrac{1}{2} \operatorname{tr}\big(p^{\mathrm{T}} J^{\mathrm{T}} \nabla f(p) X^{\mathrm{T}} J^{\mathrm{T}} p (p^{\mathrm{T}} p)^{-1}\big) && (3.14)
\end{aligned}
$$

Inserting this into our expression we end up with:

$$
\mathrm{d} f_p(X) + \tfrac{1}{2} \operatorname{tr}\big(p^{\mathrm{T}} J \nabla f(p) \underbrace{X^{\mathrm{T}} J^{\mathrm{T}} p}_{=-X^{\mathrm{T}} J p} (p^{\mathrm{T}} p)^{-1}\big) + \tfrac{1}{2} \operatorname{tr}\big(p^{\mathrm{T}} J^{\mathrm{T}} \nabla f(p) \underbrace{p^{\mathrm{T}} J^{\mathrm{T}} X}_{=-p^{\mathrm{T}} J X} (p^{\mathrm{T}} p)^{-1}\big) = \mathrm{d} f_p(X), \tag{3.15}
$$

where the last two terms cancel by the tangent space condition ref ?? $p^{\mathrm{T}} J X ? - X^{\mathrm{T}} J p$. □

—

(Christoffel symbols)

$$
\begin{aligned}
&p^{\mathrm{T}} J \nabla f(p) p^{\mathrm{T}} p + p^{\mathrm{T}} J J p (\nabla f(p))^{\mathrm{T}} J p + p^{\mathrm{T}} p (\nabla f(p))^{\mathrm{T}} J p + p^{\mathrm{T}} J^{\mathrm{T}} \nabla f(p) p^{\mathrm{T}} J^{\mathrm{T}} J p && (3.16)\\
&= -p^{\mathrm{T}} J^{\mathrm{T}} \nabla f(p) p^{\mathrm{T}} J^{\mathrm{T}} J p && (3.17)
\end{aligned}
$$

4 Conclusion

But the fact that some geniuses were laughed at does not imply that all who are laughed at are geniuses. They laughed at Columbus, they laughed at Fulton, they laughed at the Wright Brothers. But they also laughed at Bozo the Clown - Sagan (1993).

Bibliography

Ghia, U., K. N. Ghia and C. T. Shin (1982). 'High-Re Solutions for Incompressible Flow Using the Navier-Stokes Equations and a Multigrid Method'. In: *Journal of Computational Physics* 48, pp. 387–411.

NTNU, Department of Marine Technology (2020). *IMT Software Wiki - LaTeX*. URL: https://www.ntnu.no/wiki/display/imtsoftware/LaTeX (visited on 15th Sept. 2020).

Sagan, Carl (1993). *Brocas brain: reflections on the romance of science*. Presidio Press.

Appendix

A Hello World Example

```
int main {
    // This is a comment
    std::cout << "Hello World from C++!" << std::endl;
    std::cout << "I am using the default style to print this code in beautiful
    ↪  colors. Since the text is so long I have to include the 'breaklines'
    ↪  option as well" << std::endl;
    return 0;
}
```

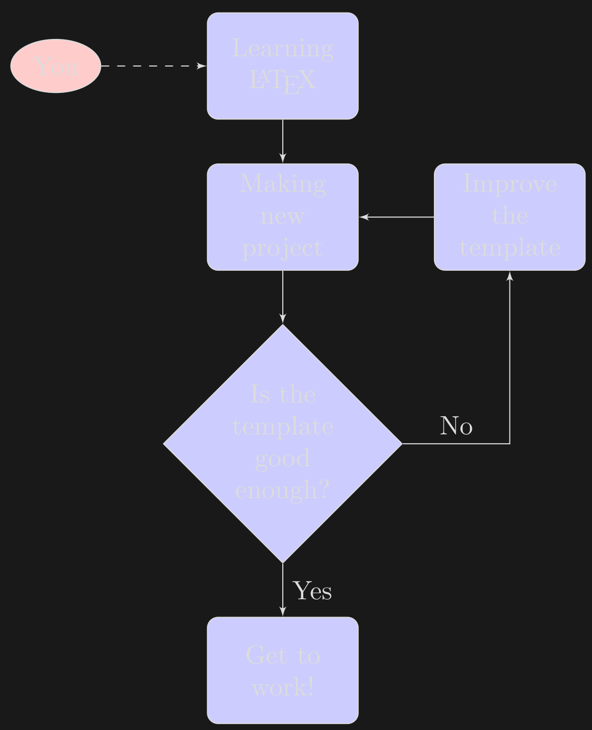
⋮

```
# This is a comment
print('Hello world from Python!')
print('I am using the "rrt" style to print this code in beautiful colors')
```

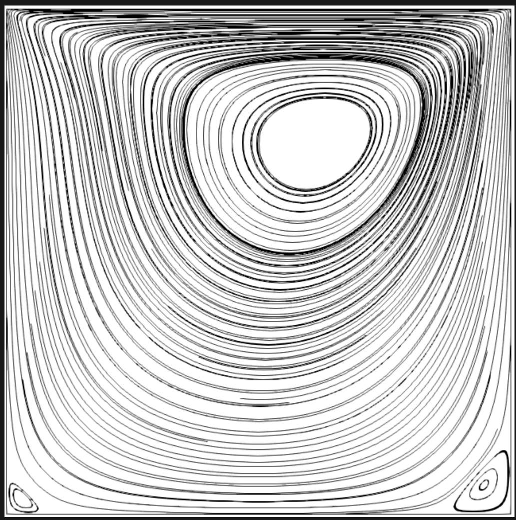
⋮

```
% This is a comment
disp("Hello World from MATLAB!");
disp("I am using the "tango" style to print this code in beautiful colors");
```

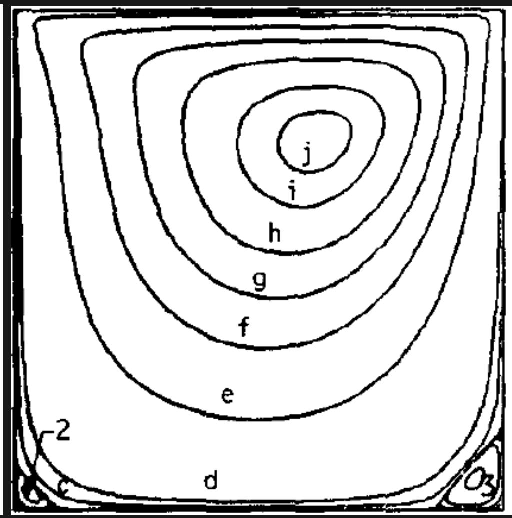
B Flow Chart Example



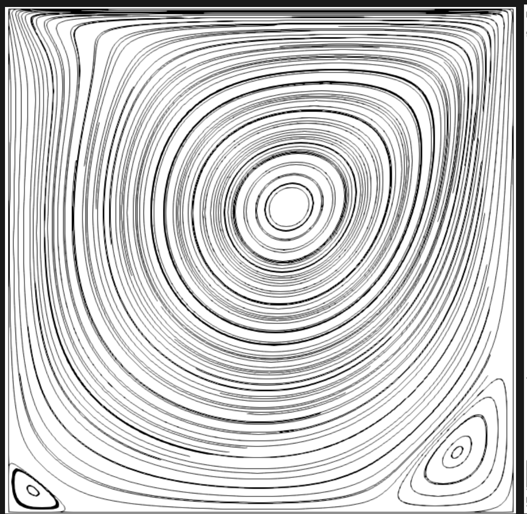
C Sub-figures Example



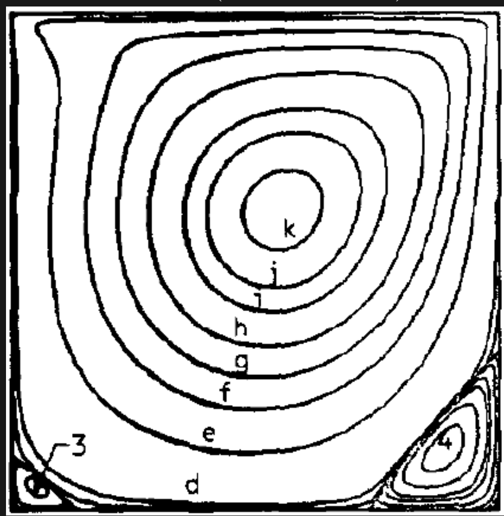
$Re = 100$ - Exercise



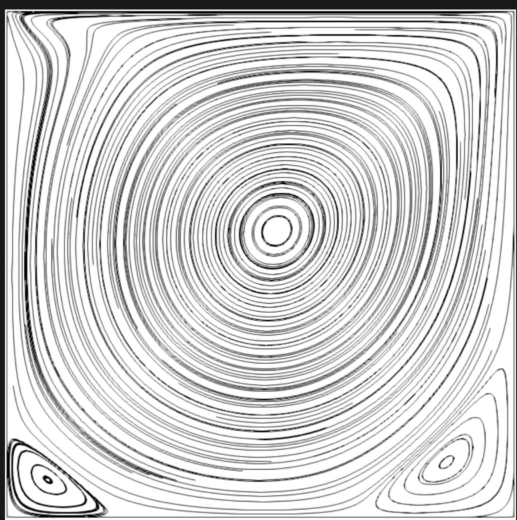
$Re = 100$ (Ghia et al. 1982)



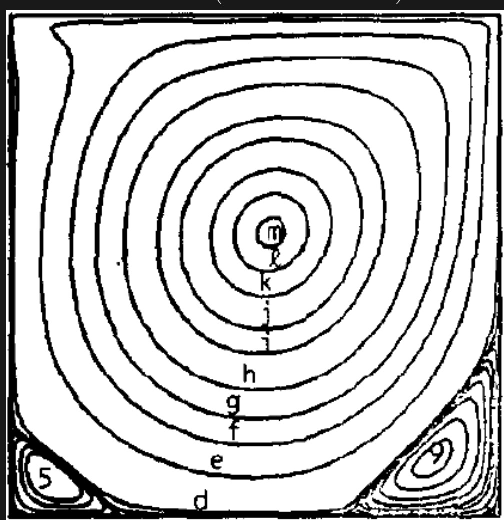
$Re = 400$ - Exercise



$Re = 400$ (Ghia et al. 1982)



$Re = 1000$ - Exercise



$Re = 1000$ (Ghia et al. 1982)

Figure 2: Streamlines for the problem of a lid-driven cavity.