Application of Kähler manifold to signal processing and Bayesian inference

Jaehyung Choi*, 1, 2
Andrew P. Mullhaupt 1

1. Department of Applied Mathematics and Statistics
2. Department of Physics and Astronomy
SUNY at Stony Brook

September 23, 2014
MaxEnt 2014, Amboise
# Table of contents

1. Review on Kähler manifold

2. Kählerian information geometry for signal processing

3. Geometric shrinkage priors

4. Example: ARFIMA

5. Conclusion
Kähler manifold and information geometry

Implications of Kähler manifold

- differential geometry, algebraic geometry
Implications of Kähler manifold

- differential geometry, algebraic geometry
- superstring theory and supergravity in theoretical physics
Kähler manifold and information geometry

Implications of Kähler manifold

- differential geometry, algebraic geometry
- superstring theory and supergravity in theoretical physics
- Information Geometry
Kähler manifold and information geometry

Implications of Kähler manifold

- differential geometry, algebraic geometry
- superstring theory and supergravity in theoretical physics
- Information Geometry
Implications of Kähler manifold

- differential geometry, algebraic geometry
- superstring theory and supergravity in theoretical physics

**Information Geometry**

- Zhang and Li (2013): symplectic and Kähler structures in divergence function
The Kähler manifold is the Hermitian manifold with the closed Kähler two-form.

In the metric expression,

\[ g_{ij} = g_{i\bar{j}} = 0 \]
\[ \partial_{i}g_{j\bar{k}} = \partial_{j}g_{i\bar{k}} = 0 \]

Any advantages? Let’s discuss later.
Linear systems and information geometry

- Linear systems are described by the transfer function $h(w; \xi)$

$$y(w) = h(w; \xi)x(w; \xi)$$

where input $x$ and output $y$.

- The metric tensor for the filter

$$g_{\mu\nu}(\xi) = \frac{1}{2\pi} \int_{-\pi}^{\pi} (\partial_{\mu} \log S)(\partial_{\nu} \log S)dw$$

where $S(w; \xi) = |h(w; \xi)|^2$. 
• **z-transformation** \( h(z; \xi) = \sum_{r=0}^{\infty} h_r(\xi)z^{-r} \)

\[
\log h(z; \xi) = \log h_0 + \log \left(1 + \sum_{r=1}^{\infty} \frac{h_r}{h_0} z^{-r}\right) = \log h_0 + \sum_{r=1}^{\infty} \eta_r z^{-r}
\]

• The metric tensor in terms of transfer function

\[
g_{\mu\nu} = \frac{1}{2\pi i} \oint_{|z|=1} \partial_{\mu} \left( \log h + \log \bar{h} \right) \partial_{\nu} \left( \log h + \log \bar{h} \right) \frac{dz}{z}
\]

where \( \mu, \nu \) run holomorphic and anti-holomorphic indices.
The metric tensors in holomorphic and anti-holomorphic coordinates

\[ g_{ij}(\xi) = \frac{1}{2\pi i} \oint_{|z|=1} \partial_i \log h(z; \xi) \partial_j \log h(z; \xi) \frac{dz}{z} \]

\[ g_{i\bar{j}}(\xi) = \frac{1}{2\pi i} \oint_{|z|=1} \partial_i \log h(z; \xi) \partial_{\bar{j}} \log \bar{h}(\bar{z}; \bar{\xi}) \frac{dz}{z} \]

The metric tensor

\[ g_{ij} = \partial_i \log h_0 \partial_j \log h_0 \]

\[ g_{i\bar{j}} = \partial_i \log h_0 \partial_{\bar{j}} \log \bar{h}_0 + \sum_{r=1}^{\infty} \partial_i \eta_r \partial_{\bar{j}} \bar{\eta}_r \]
Kähler manifold for signal processing

**Theorem**

Given a holomorphic transfer function $h(z; \xi)$, the information geometry of a signal processing model is Kähler manifold if and only if $h_0$ is a constant in $\xi$.

$(\Rightarrow)$ If the geometry is Kähler, it should be Hermitian imposing

$$g_{ij} = \partial_i \log(h_0) \partial_j \log(h_0) = 0 \rightarrow h_0 \text{ constant in } \xi$$

$(\Leftarrow)$ If $h_0$ is a constant in $\xi$, the metric tensor is given in

$$g_{ij} = 0 \text{ and } g_{i\bar{j}} = \sum_{r=1}^{\infty} \partial_i \eta_r \partial_{\bar{j}} \bar{\eta}_r \rightarrow \text{Hermitian}$$

The Kähler two-form is closed: $\Omega = ig_{i\bar{j}} d\xi^i \wedge d\bar{\xi}^j$
Kähler potential for signal processing

On the Kähler manifold, the metric tensor is

$$g_{ij} = \partial_i \partial_j \mathcal{K}$$

where the Kähler potential $\mathcal{K}$.

**Corollary**

Given Kähler geometry, the Kähler potential of the geometry is the square of the Hardy norm of the log-transfer function.

$$\mathcal{K} = \frac{1}{2\pi i} \int_{|z|=1} (\log h(z; \xi))(\log h(z; \xi))^* \frac{dz}{z}$$

$$= \| \log h(z; \xi) \|^2_{H^2}$$
Benefits of Kählerian information geometry

1. Calculation of geometric objects is simplified.

\[ g_{i\bar{j}} = \partial_i \partial_{\bar{j}} \mathcal{K}, \quad \Gamma_{i\bar{j},\bar{k}} = \partial_i \partial_{\bar{j}} \partial_{\bar{k}} \mathcal{K} \]

\[ R^i_{j\bar{m}n} = \partial_{\bar{m}} \Gamma^i_{jn}, \quad R_{i\bar{j}} = -\partial_i \partial_{\bar{j}} \log \mathcal{G} \]

2. Easy $\alpha$-generalization and linear order correction in $\alpha$

\[ \Gamma^{(\alpha)} = \Gamma + \alpha T, \quad R^{(\alpha)} = R + \alpha \partial T \]

3. Submanifolds of Kähler is Kähler.

4. Laplace-Beltrami operator: $\Delta = 2g^{i\bar{j}} \partial_i \partial_{\bar{j}}$
Komaki (2006): The difference in risk functions is given by

\[
\mathbb{E}(D_{KL}(p(y|\xi)||p_{\pi_J}(y|x^{(N)}))|\xi)) - \mathbb{E}(D_{KL}(p(y|\xi)||p_{\pi_I}(y|x^{(N)}))|\xi)) \\
= \frac{1}{2N^2} g^{ij} \partial_i \log \left( \frac{\pi_I}{\pi_J} \right) \partial_j \log \left( \frac{\pi_I}{\pi_J} \right) - \frac{1}{N^2} \frac{\pi_J}{\pi_I} \Delta \left( \frac{\pi_I}{\pi_J} \right) + o(N^{-2})
\]

If \( \psi = \frac{\pi_I}{\pi_J} \) is superharmonic, \( p_{\pi_I} \) outperforms \( p_{\pi_J} \).
Superharmonic prior \( \pi_I \), Jeffreys prior \( \pi_J \)
Superharmonicity of functions is hard to check.
In particular, in high-dimensional curved geometry!
On a Kähler manifold, a positive function \( \psi = \Psi(u^* - \kappa(\xi, \bar{\xi})) \) is a superharmonic prior function if \( \kappa(\xi, \bar{\xi}) \) is (sub)harmonic, bounded above by \( u^* \), and \( \Psi \) is concave decreasing: \( \Psi'(\tau) > 0, \Psi''(\tau) < 0 \).

The ansätze for \( \Psi \):

\[
\Psi_1(\tau) = \tau^a, \quad \Psi_2(\tau) = \log(1 + \tau^a) \quad (\tau > 0, 0 < a \leq 1)
\]

The ansätze for \( \kappa \):

\[
\kappa_1 = \mathcal{K}, \kappa_2 = \sum_{r=0}^{\infty} a_r |h_r(\xi)|^2, \kappa_3 = \sum_{i=1}^{n} b_i |\xi^i|^2 \quad (a_r > 0, b_i > 0)
\]
Algorithm for geometric priors

The algorithm for finding geometric priors is the following:

1. Check whether the geometry is Kähler.
2. Check the superharmonicity of prior function $\psi$.
3. If (sub)harmonic, plug it into the theorem to get superharmonic functions and move to the next step.
4. If superharmonic, multiply the Jeffreys prior and set it as the shrinkage prior.
5. Do Bayesian inference.
The transfer function of ARFIMA:

\[ h(z; \xi) = \frac{(1 - \mu_1 z^{-1})(1 - \mu_2 z^{-1}) \cdots (1 - \mu_q z^{-1})}{(1 - \lambda_1 z^{-1})(1 - \lambda_2 z^{-1}) \cdots (1 - \lambda_p z^{-1})} (1 - z^{-1})^d \]

The Kähler potential:

\[ \mathcal{K} = \sum_{n=1}^{\infty} \left| \frac{d + (\mu_1^n + \cdots + \mu_q^n) - (\lambda_1^n + \cdots + \lambda_p^n)}{n} \right|^2 \]

The metric tensor of ARFIMA:

\[
g_{\bar{i}j} = \begin{pmatrix}
\frac{\pi^2}{6} & \frac{1}{\lambda_j} \log (1 - \bar{\lambda}_j) & -\frac{1}{\bar{\mu}_j} \log (1 - \bar{\mu}_j) \\
\frac{1}{\lambda_i} \log (1 - \lambda_i) & \frac{1}{1 - \lambda_i \lambda_j} & -\frac{1}{1 - \lambda_i \bar{\mu}_j} \\
-\frac{1}{\mu_i} \log (1 - \mu_i) & -\frac{1}{1 - \mu_i \lambda_j} & \frac{1}{1 - \mu_i \bar{\mu}_j}
\end{pmatrix} \]
Conclusion

- Kähler manifold: information geometry for signal processing
- Kähler potential: square of Hardy norm of log-transfer function
- Several computational benefits exist on the Kähler manifold.
- In particular, Komaki priors are easy to build.
- An algorithm and ansätze for Komaki priors are introduced.


Tanaka, F., Superharmonic priors for autoregressive models, Mathematical Engineering Technical Reports, University of Tokyo (2009)

Zhang, J. and Li, F., Symplectic and Kähler Structures on Statistical Manifolds Induced from Divergence Functions, Geometric Science of Information 8085 (2013) 595-603