Matrix realization of homogeneous Hessian domains

Hideyuki ISHI
(Nagoya University/ JST PRESTO)

3rd conference on Geometric Science of Information
November 7–9, 2017, Mines ParisTech
§1. Homogeneous Hessian domains
§2. Main result
§3. Realization of 2-dimensional Hessian domains

The formulation is slightly different from the proceedings.
1. Homogeneous Hessian domains

\( \mathbb{R}^n \supset D : \) domain

\( g : \) Hessian metric on \( D \), i.e.

\[ \forall x \in D \ \exists \text{nbhd } U \ni x \ \exists \varphi_U \in C^\infty(U) \text{ s.t.} \]

\[ g|_U = \sum_{i,j=1}^{n} \frac{\partial^2 \varphi_U}{\partial x_i \partial x_j} dx_i dx_j. \]

Such a Hessian domain \((D, g)\) often appears in Information Geometry, e.g. as the parameter space of exponential family (e-coordinate with the Fisher metric).
\[(D, g) \cong (D', g') : \text{isomorphic}\]
\[\iff \exists \Phi : \mathbb{R}^n \to \mathbb{R}^n \text{ (affine)} \st\ \Phi(D) = D' \text{ and } \Phi^*g' = g.\]

The automorphism group of the Hessian domain \((D, g)\) is defined by
\[
\text{Aut}(D, g) := \{ \Phi \in \text{Aff}(\mathbb{R}^n) \mid \Phi(D) = D, \ \Phi^*g = g \}.\]

\((D, g) : \text{homogeneous}\)
\[\iff \text{The group Aut}(D, g) \text{ acts on } D \text{ transitively, i.e.}\]
\[\forall x, y \in \text{Aut}(D, g) \ \exists \Phi \in \text{Aut}(D, g) \st y = \Phi(x).\]

A Basic theory for homogeneous Hessian domains is established by H. Shima (1980).

Theorem (Shima) Every homogeneous Hessian domain is convex.
Example 1 (The Euclidean space)
\[ \mathcal{D} = \mathbb{R}^n, \quad g = \sum dx_i^2. \]
\[ \varphi(x) := \sum (x_i)^2/2 = \langle x, x \rangle/2 ; \text{ potential of } g. \]

The metric \( g \) is invariant under the translation \( x \mapsto x + a \) by any \( a \in \mathbb{R}^n \).
Indeed, we have
\[ \varphi(x + a) = \varphi(x) + \langle a, x \rangle + \langle a, a \rangle/2. \]

The parameter space of the normal distributions \( \{ N(x, I) \}_{x \in \mathbb{R}^n} \) with the identity covariance matrix is naturally identified with this Hessian domain.
Example 2 (Positive definite cone)
\[ \Omega_n = \{ X \in \text{Sym}(n, \mathbb{R}) \mid X \text{ is positive definite} \} \]
\[ \subset \text{Sym}(n, \mathbb{R}) \equiv \mathbb{R}^{n(n+1)/2}. \]
For \( s > 0 \), define
\[ g_s(U, U')_X := s \text{tr}(X^{-1}UX^{-1}U') \quad (X \in \Omega_n, U, U' \in \text{Sym}(n, \mathbb{R})). \]
\[ \varphi_s(X) := -s \log \det X \quad (X \in \Omega_n) ; \text{ potential of } g_s. \]

The metric \( g_s \) is \textit{invariant under} \( \rho(A) : X \mapsto AX^tA \) with any \( A \in GL(n, \mathbb{R}) \). Indeed, we have
\[ \varphi_s(\rho(A)X) = \varphi_s(X) - 2s \log \det A. \]
The group $\text{Aut}(\Omega_n, g_s) = \{ \rho(A) | A \in GL(n, \mathbb{R}) \}$ acts on $\Omega_n$ transitively. Indeed, for $X, Y \in \Omega_n$, taking the Cholesky decompositions $X = T^tT$, $Y = S^tS$, we have $Y = \rho(A)X$ with $A := ST^{-1}$.

The parameter space of the central normal distributions $\{N(0, \Sigma)\}_{\Sigma \in \Omega_n}$ is naturally identified $(\Omega_n, g_{1/2})$ via $X = \Sigma^{-1/2}$. 
For \( s = (s_1, \ldots, s_n) \in \mathbb{R}_0^n \), define

\[
\varphi_s(X) := -s_n \log \det X + \sum_{k=1}^{n-1} (s_k - s_{k+1}) \log \det X^{[k]},
\]

for \( X \in \Omega_n \), where \( X^{[k]} := (X_{ij})_{i,j \leq k} \in \text{Sym}(k, \mathbb{R}) \) and let \( g_s \) be the Hessian of \( \varphi_s \). For a scalar \( s > 0 \), we have \( g_s = g(s,s,\ldots,s) \).

Let \( H_n \) be the group of lower triangular matrices with positive diagonals. Then \( \{ \rho(A) \mid A \in H_n \} \subset \text{Aut}(\Omega_n, g_s) \), which acts on \( \Omega_n \) transitively.

Every homogeneous Hessian domain on \( \Omega_n \) is isomorphic to \( g_s \) with a unique \( s \in \mathbb{R}_0^n \).
Example 3
\[ \mathcal{U}_n := \Omega_n \times \mathbb{R}^n \subset \mathbb{R}^{n(n+1)/2+n}. \]
For \( s > 0 \), define
\[
g_s((U, v), (U', v'))_{(X, w)} = s \text{tr}(X^{-1}UX^{-1}U') + twX^{-1}UX^{-1}U'X^{-1}w - 2twX^{-1}(UX^{-1}v' + U'X^{-1}v) + 2tvX^{-1}v' \]
\[
((X, w) \in \mathcal{U}_n, (U, v), (U', v') \in \text{Sym}(n, \mathbb{R}) \oplus \mathbb{R}^n) ,
\]
\[ \varphi_s(X, w) := -s \log \det X + twX^{-1}w \quad ((X, w) \in \mathcal{U}_n). \]
The metric \( g_s \) is invariant under the following:
\[
\sigma(A) : (X, w) \mapsto (AX^tA, Aw) \quad (A \in GL(n, \mathbb{R})) ,
\]
\[
\tau(b) : (X, w) \mapsto (X, w + Xb) \quad (b \in \mathbb{R}^n) .
\]
The group \( \text{Aut}(\mathcal{U}_n, g_s) \) acts on \( \mathcal{U}_n \) transitively.

The parameter space of \( \{N(\mu, \Sigma)\}_{\mu \in \mathbb{R}^n, \Sigma \in \Omega_n} \) is naturally identified with \( (\mathcal{U}_n, g_{1/2}) \) via \( X = \Sigma^{-1}/2, w = \Sigma^{-1}\eta \).
\[ V_n := \text{Sym}(n + 1, \mathbb{R})/\mathbb{R}E_{n+1,n+1} \]
\[ = \left\{ \begin{pmatrix} X & w \\ tw & * \end{pmatrix} \mid X \in \text{Sym}(n, \mathbb{R}), w \in \mathbb{R}^n \right\}. \]

Then
\[ \rho \begin{pmatrix} A & 1 \\ tw & * \end{pmatrix} \begin{pmatrix} X & w \\ tw & * \end{pmatrix} = \begin{pmatrix} AX^tA & Aw \\ tw^tA & * \end{pmatrix}, \]
\[ \rho \begin{pmatrix} I & 1 \\ tb & 1 \end{pmatrix} \begin{pmatrix} X & w \\ tw & * \end{pmatrix} = \begin{pmatrix} X & w + Xb \\ t(w + Xb) & * \end{pmatrix}. \]

For \( s = (s_1, \ldots, s_n) \in \mathbb{R}^n_{>0} \), define
\[ \varphi_s(X, w) := -s_n \log \det X + \sum_{k=1}^{n-1} (s_k - s_{k+1}) \log \det X^k \\
+ t_w X^{-1} w \quad ((X, w) \in \mathcal{U}_n) \]
and let \( g_s \) be the Hessian of \( \varphi_s \). Then \( g_s \) is invariant under \( \sigma(A) \) with \( A \in H_n \) and \( \tau(b) \) (\( b \in \mathbb{R}^n \)).
Remark: As a Hessian domain, \((\mathcal{U}_n, g_s)\) is NOT isomorphic to the direct product of Example 1 (Euclidean space) and Example 2 (Positive definite cone).
§2. Main result
Every homogeneous Hessian domain is isomorphic to
\((U_N \cap \Gamma, g_s|U_N \cap \Gamma)\) with an appropriate affine subspace \(\Gamma\)
of \(V_N\) and \(s \in \mathbb{R}^N_{>0}\), where \(N\) can be taken to be less
than or equal to the dimension of the domain.

See the proceedings for a construction of \(\Gamma\) as a set
of matrices with specific block decomposition.

Application (in Future): Realization of homogeneous
exponential families.
§3. Realization of 2-dimensional Hessian domains

(I) The Euclidean space

\[ \mathbb{R}^2 \ni (w_1, w_2) \mapsto \begin{pmatrix} 1 & 0 & w_1/\sqrt{2} \\ 0 & 1 & w_2/\sqrt{2} \\ w_1/\sqrt{2} & w_2/\sqrt{2} & * \end{pmatrix} \in \mathcal{U}_2. \]

(II) The direct product of half-liness

\[ \mathbb{R}_{>0}^2 \ni (x_1, x_2) \mapsto \begin{pmatrix} x_1 & 0 & 0 \\ 0 & x_2 & 0 \\ 0 & 0 & * \end{pmatrix} \in \mathcal{U}_2. \]

(III) The direct product \( \mathbb{R} \times \mathbb{R}_{>0} \)

\[ \mathbb{R} \times \mathbb{R}_{>0} \ni (w, x) \mapsto \begin{pmatrix} 1 & 0 & w/\sqrt{2} \\ 0 & x & 0 \\ w/\sqrt{2} & 0 & * \end{pmatrix} \in \mathcal{U}_2. \]
(IV) The domain $\mathcal{U}_1$

The metric $g_s \ (s > 0)$ on $\mathcal{U}_1 = \left\{ \begin{pmatrix} x & w \\ w & * \end{pmatrix} \mid x > 0, w \in \mathbb{R} \right\}$ is written by

$$g_s(x, w) = \left( \frac{sx + 2w^2}{x^3} \right) dx^2 - \frac{4w}{x^2} dx dw + \frac{2}{x} dw^2.$$
(V) Parabolic domain
\[ \mathcal{D} = \left\{ (x, y) \mid y > x^2 \right\} \subset \mathbb{R}^2 \]
For \( s > 0 \), define
\[ g_s(x, y) = \frac{s}{(y-x^2)^2} \{(2y + 2x^2)dx^2 - 2xdxdy + dy^2\}, \]
\[ \varphi_s(x, y) := -s \log(y - x^2) \quad ((x, y) \in \mathcal{D}) \]

The metric \( g_s \) is invariant under the following:
\( (x, y) \mapsto (cx, c^2y) \) \( (c \neq 0) \)
\( (x, y) \mapsto (x + b, y + 2bx + b^2) \) \( (b \in \mathbb{R}) \).

The group \( \text{Aut}(\mathcal{D}, g_s) \) is generated by the these transforms, and acts on \( \mathcal{D} \) transitively.
We have the imbedding

\[ D \ni (x, y) \mapsto \begin{pmatrix} 1 & x & 0 \\ x & y & 0 \\ 0 & 0 & * \end{pmatrix} \in U_2. \]

\[
\rho \begin{pmatrix} 1 & 0 & 0 \\ 0 & c & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & x & 0 \\ x & y & 0 \\ 0 & 0 & * \end{pmatrix} = \begin{pmatrix} 1 & cx & 0 \\ cx & c^2y & 0 \\ 0 & 0 & * \end{pmatrix},
\]

\[
\rho \begin{pmatrix} 1 & 0 & 0 \\ b & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & x & 0 \\ x & y & 0 \\ 0 & 0 & * \end{pmatrix} = \begin{pmatrix} 1 & x + b & 0 \\ x + b & y + 2bx + b^2 & 0 \\ 0 & 0 & * \end{pmatrix}.
\]

This domain \( D \) and \( U_1 \) are mutually dual. See

\[ U_1 \cong \left\{ \begin{pmatrix} 1 & 0 & 0 \\ 0 & x & w \\ 0 & w & * \end{pmatrix} \mid x > 0, w \in \mathbb{R} \right\}. \]