

Visualization and Inference From a Point Cloud on a Curved, Singular Manifold*

A. Ronald Gallant[†]

First draft: August 30, 2022

This draft: August 16, 2023

Paper: www.aronaldg.org/papers/curve.pdf

Code: www.aronaldg.org/webfiles/curve_ds_eg

*Address correspondence to A. Ronald Gallant, P.O. Box 659, Chapel Hill NC 27514, USA, phone 919-428-1130; email aronldg@gmail.com.

[†]Emeritus Professor of Economics, University of North Carolina; Emeritus Professor of Economics, Penn State University.

Abstract

We consider credibility regions computed from a point cloud consisting of draws from a posterior that lie on a singular manifold that is embedded in a natural Euclidean parameter space. Visualization methods are developed to determine the amount of curvature of the manifold. Methods to visualize and report credibility regions in the presence of curvature are proposed. The motivating application is MCMC (Markov Chain Monte Carlo) applied to a likelihood that is subject to overidentified moment equations. A common approach when analyzing such data is to map the data to an Euclidean space of the same dimension as the manifold, called a chart, with distance on the chart equal to geodesic distance on the manifold. That approach is adopted here with the difference that our chart variables are interpretable. Among the examples is a replication of the classic Hansen and Singleton (1982) estimation using their original data and the methods proposed here.

Keywords and Phrases: Credibility regions, Point cloud, Curved, singular manifold, Bayesian inference, Method of moments.

JEL Classification: C11, C14, C15, C32, C36, C58

1 Introduction

Constructing a $(1 - \alpha) \times 100\%$ credibility region from posterior draws is straightforward: One centers a geometric shape at a reasonable point in the parameter space and adjusts the boundary until $(1 - \alpha) \times 100\%$ of the draws are within the region. If the draws lie in a curved manifold that is both contained within a higher dimensional, natural parameter space and singular with respect to Lebesgue measure on the natural parameter space, then the credibility region is also a curved, singular manifold. The problem considered here is how to visualize and report that credibility region with respect to the natural parameter space.

A standard approach when analyzing such data is to map the data to a Euclidean space of the same dimension as the manifold, called a chart, with distance on the chart equal to geodesic distance on the manifold. That approach is adopted here with the difference from the standard approach in that chart variables are a subset of parameter space variables.

The situation just described can arise as follows:¹ A likelihood

$$f(y | x, \rho) = \prod_{t=1}^n f(y_t | x_{t-1}, \rho), \quad (1)$$

is available where y_t is a column vector and x_{t-1} is a matrix of exogenous and predetermined variables with a fixed number of rows. The vector ρ contains the location, scale, etc. parameters of the likelihood. The number of columns of x_{t-1} is either fixed, as in a cross-sectional model or a VAR model or increasing with t , as in a VAR-GARCH model.² The y and x are objects that contain the observed y_t and x_{t-1} . The likelihood can be a sieve with variable number of parameters thus making the Bayes estimator nonparametric.

Estimation of the parameters in (1) is subject to moment conditions³

$$0 = q(\rho, \theta) = \frac{1}{n} \sum_{t=1}^n \int \mathbf{m}(y, x_{t-1}, \rho, \theta) f(y | x_{t-1}, \rho) dy, \quad q \in \mathbb{R}^m \quad (2)$$

support conditions

$$h(\rho, \theta) > 0, \quad h \in \mathbb{R}^l \quad (3)$$

¹This description follows Gallant (2022a) with slight modification.

²This is due to the recursive structure of GARCH variance which causes a VAR-GARCH model to be non-Markovian and to depend on the past up to the initial observation as most VAR-GARCH likelihoods are implemented in practice.

³Note that \mathbf{m} and m in (2) are distinguished. One can integrate with respect to the distribution of x_{t-1} rather than the empirical distribution of x_{t-1} if it is available.

and a prior

$$\pi(\rho, \theta). \tag{4}$$

The vector θ contains the parameters of the scientific model that determine the moment conditions and that are not in ρ . The natural parameter space is \mathbb{R}^{d_a} where d_a is the sum of the dimensions of ρ and θ . Distance between two points in the natural parameter space is measured along a straight line using the Euclidean norm.

Letting⁴ $\mathbf{x} = (\rho, \theta)$, the support of the posterior density $p(\rho, \theta | x, y)$ is the manifold

$$M = \{\mathbf{x} \in \mathbb{R}^{d_a} : q_i(\mathbf{x}) = 0, i = 1, \dots, m, h_j(\mathbf{x}) > 0, j = 1, \dots, l\} \tag{5}$$

The parameters ρ are induced in $q(\rho, \theta)$ by the integration. They may also appear explicitly in $\mathbf{m}(y_t, x_{t-1}, \rho, \theta)$ as the notation indicates but in most applications \mathbf{m} does not actually depend on ρ .

We assume overidentification, i.e., that the dimension m of $q(\rho, \theta)$ is larger than the dimension of θ . Under this setup, the support of the posterior density $p(\rho, \theta | x, y)$ is singular with respect to Lebesgue measure on \mathbb{R}^{d_a} (Bornn, Shephard, and Solgi, 2018). Let $d = d_a - m$ denote the dimension of the support.

The posterior density $p(\rho, \theta | x, y)$ is determined by (1) through (3). It is not known in closed form but can be simulated (Gallant, 2022a)..

Often $f(y|x, \rho)$ is a sieve so that the elements of the parameter ρ determine location, scale, heteroskedasticity, etc. An interesting attempt to make both $f(y|x, \rho)$ and $\mathbf{m}(y, x_{t-1}, \rho, \theta)$ nonparametric is Gallant and Tauchen (1989).

The example in Subsection 7.4 is an illustration of the above considerations.

The point cloud need not be generated in this way. It could be a simulation other than by MCMC from an analytically intractable posterior. Or generated by a scientific apparatus. The relevant characteristics of the motivating problem are that d , $q(\rho, \theta)$, and $p(\rho, \theta | x, y)$ are known.⁵ Here, we shall presume that $q(\rho, \theta)$ and $p(\rho, \theta | x, y)$ are not known.

⁴In this paper, sans serif \mathbf{x} and \mathbf{y} are distinguished from italic x and y ; the former referring to parameters and the later to data. This is to maintain compatibility with both econometric conventions and the numerical analysis conventions of Zappa, Holmes-Cerfon, and Goodman (2018) and Gallant (2022a, 2022b).

⁵For the motivating problem, the normalizing constant of $p(\rho, \theta | x, y)$ is compatible (Gallant, 2022a, Subsection 2.4) although knowledge to within a multiplicative constant is usually adequate.

Denote the point cloud by

$$\mathcal{D} = \{\mathbf{x}_i\}_{i=1}^N = \{(\rho_i, \theta_i)\}_{i=1}^N. \quad (6)$$

For a set R that may be either a subset of the manifold M or of the embedding parameter space \mathbb{R}^{d_a} , denote probability with respect to the empirical distribution determined by \mathcal{D} as

$$P_{\mathcal{D}}(R | x, y) = \frac{1}{N} \sum_{i=1}^N I(\mathbf{x}_i \in R). \quad (7)$$

On the manifold M distance is computed along geodesics. One computes distance by traversing a geodesic from a starting point s to an end point p and accumulating (infinitesimal increments of) a Hausdorff weight function defined on M (Morgan, 2016). Denote geodesic distance between two points s and p in M by

$$\delta_G(s, p) \quad (8)$$

and denote Euclidean distance between them by

$$\delta_E(s, p). \quad (9)$$

A measure of the location of the point cloud is the intrinsic mean $\bar{\mathbf{x}}$. It is the point s in \mathcal{D} that minimizes $\frac{1}{N} \sum_{i=1}^N \delta_G(s, \mathbf{x}_i)$.

Letting

$$V = V_{IC} = \frac{1}{N} \sum_{i=1}^N (\mathbf{x}_i - \bar{\mathbf{x}})(\mathbf{x}_i - \bar{\mathbf{x}})^{\top}, \quad (10)$$

we give primary attention to two regions. The first is a rectangle

$$R_r = \times_{k=1}^{d_a} [\bar{x}_k - \tau \sqrt{v_{kk}}, \bar{x}_k + \tau \sqrt{v_{kk}}], \quad (11)$$

where \bar{x}_k denotes an element of $\bar{\mathbf{x}}$, and the v_{kk} are the diagonal elements of V . Choosing τ such that $P_{\mathcal{D}}(R_r | x, y) = 1 - \alpha$ gives a set of simultaneously valid $(1 - \alpha) \times 100\%$ credibility intervals. An advantage of this region is that it is amenable to tabular reporting.

The second is an ellipse

$$R_e = \{\mathbf{x} \in \mathcal{D} \mid (\mathbf{x} - \bar{\mathbf{x}})^{\top} V^{-1} (\mathbf{x} - \bar{\mathbf{x}}) < \tau\}, \quad (12)$$

with τ such that $P_{\mathcal{D}}(R_e | x, y) = 1 - \alpha$. Aside from being a traditional region, it can be used to localize the point cloud around \bar{x} to facilitate certain computations.

Our interest is in draws from exact data generating mechanisms. That is, the draws $\mathcal{D} = \{\mathbf{x}_i\}_{i=1}^N$ are in M to within reasonable precision for linear algebra on a machine. To our knowledge, there are three algorithms that can satisfy this requirement for the Bayesian inference problem defined by (1) through (4). Gallant (2022a) generates draws $\{\mathbf{x}_i\}_{i=1}^N$ in M for the problem as stated by using the Surface Sampling Algorithm of Zappa, Holmes-Cerfon, and Goodman (2018). Bornn, Shephard, and Solgi require that (1) has discrete support, which makes (2) a sum involving probability weights and their corresponding support. Their paper contains numerous examples and an extensive review of literature related to this problem. Shin presumes that (1) is a mixture of specific parametric distributions with random weights drawn from a discrete distribution. The constraint (2) becomes a constraint on the discrete distribution of the random weights. His examples are from macro economics.

In what follows, requirements are as follows: The manifold M that contains the point cloud \mathcal{D} must be connected and have the same dimension d everywhere. Either the point cloud does not contain duplicate points or they are easily detected. If the point cloud is an MCMC chain then duplicate points must occur in succession and therefore are easy to detect.

Code, including a User’s Guide, implementing methods introduced here for the SNP sieve $f(y_t | x_{t-1}, \rho)$ proposed by Gallant and Nychka (1987) as adapted to time series applications by Gallant and Tauchen (1989) is at <http://www.aronaldg.org/webfiles/npb>.

Parts of this paper borrow from Gallant (2022a, 2022b) so as to make this paper mostly self contained.

2 Geodesics

An approximate geodesic between two points s and p on M can be computed by centering ϵ -balls at each point in the cloud that are just large enough that their union M_ϵ is a connected subset of \mathbb{R}^{d_a} . The approximate geodesic is the path with shortest Euclidean distance along line segments between s and p that stay within M_ϵ . The approximation improves as ball radius becomes smaller due to adding points to the cloud (Memoli and Sapiro, 2001).

The Fast Marching Algorithm of Sethian (1996) is a standard method for finding the shortest path between a selected point s and all other points in M_ϵ but, unfortunately, it is limited to dimensions d_a less than about five due the demands on computer memory caused by to having to interpolate the point cloud to an equally spaced grid. Regardless of dimension, the method described next is far more convenient for a point cloud.

Rather than an interpolated, equally spaced grid, one can let the point cloud determine an unequally spaced grid and use Dijkstra’s algorithm (Dijkstra, 1959) to compute geodesics. If M_ϵ is a connected set, then the points in \mathcal{D} may be viewed as nodes p_j of a graph \mathcal{G}_ϵ connected by edges e_{j,j^\top} that have Euclidean length $\delta_E(p_j, p_{j^\top})$ and that stay within M_ϵ . In view of the fact that $\mathbf{x}_i \in \mathcal{D}$ are draws from a posterior and the contours of the posterior density are not spheres, our ϵ -balls for determining \mathcal{G}_ϵ are rectangles with sides k equal to $\Delta \max\{|\mathbf{x}_{k,i} - \mathbf{x}_{k,i-1}| : \mathbf{x}_i \in \mathcal{D}\}$ where $\mathbf{x}_{k,i}$ denotes the k th element of \mathbf{x}_i . If the point cloud \mathcal{D} is an MCMC chain, then \mathcal{D} will contain duplicates due to rejections. They are easily detected because they must occur in succession. In this case, nodes are the distinct points $\mathcal{D}^* = \{p_j\}_{j=1}^{N^*}$; $j(i)$ is the mapping from the point cloud index i to the node index j , which is the identity map if there are no duplicates. This construction is necessitated by the requirements of Dijkstra’s algorithm. As will be seen in the remainder, all of the points in \mathcal{D} are used in other computations whether duplicates or not by means of the mapping $j(i)$. See, e.g, (14).

From a start s , Dijkstra’s algorithm returns the shortest path that traverses edges to every node p_j and returns the path $(j_1^p, j_2^p, \dots, j_k^p)$ that connects them, where j_1^p refers to starting node s and j_k^p to ending node p_j . The approximation to $\delta_G(s, p_j)$ is

$$\delta_G(s, p_j) \approx \sum_{i=2}^k \delta_E(p_{j_{i-1}^p}, p_{j_i^p}). \quad (13)$$

Computations are as follows.

One proceeds by choosing a Δ and constructing the graph \mathcal{G}_ϵ . If \mathcal{G}_ϵ is not connected, Dijkstra’s algorithm will return ∞ for the distance from s to an isolated node. As Δ decreases the number of isolated nodes increases. At first these isolated nodes typically occur on the fringes of the point cloud and are few in number. Again, typically, as Δ decreases further, the number of isolated nodes increases abruptly. The recommendation here is to choose

Δ such that the number of isolated nodes is either few or none and to remove the points corresponding to unreachable nodes from the point cloud. This leaves a point cloud where all nodes are reachable for that Δ .

Note that the presence of spurious edges in \mathcal{G}_ϵ increases as Δ increases. Spurious edges cause the difference between δ_G and δ_E to decrease. Indeed, for Δ large enough, all nodes become connected and δ_G becomes δ_E . The method for detecting curvature proposed later becomes less sensitive as Δ increases because it depends on the difference between δ_G and δ_E . This suggests that for the purpose of detecting curvature, it is better to lean toward small Δ . Hence the advice for choosing Δ above.

Hereafter, Δ is fixed and \mathcal{D} and $P_{\mathcal{D}}$ given by (6) and (7) refer to a cloud for which all points are reachable for that Δ .

3 The Intrinsic Mean and the Chart

As mentioned previously, the estimate of the intrinsic mean \bar{x} is the start s for Dijkstra’s algorithm that minimizes the average distance to all points in the point cloud. I.e.

$$\bar{x} = \operatorname{argmin}_{s \in \mathcal{D}} \frac{1}{N} \sum_{i=1}^N \delta_G^2(s, p_{j(i)}), \quad (14)$$

where nodes $\mathcal{D}^* = \{p_j\}_{j=1}^{N^*}$ are the distinct points in the point cloud $\mathcal{D} = \{\mathbf{x}_i\}_{i=1}^N$ and $j(i)$ maps the point cloud index i to the node index j .

Computing the intrinsic mean is an order N^2 computation and can be quite time consuming. One way to reduce run times is to search only among likely candidates for the mean. One can, say, randomly divide \mathcal{D} into ten subsets and unrestrictedly search for the intrinsic mean in the first subset. Then search only among, say, the thousand nodes closest to the mean found in the first subset in the union of the first and second. Continue so on until one is searching among, say, the closest one hundred in the full cloud. A better suggestion might be to start the computation searching all of \mathcal{D} then go out for coffee.

A common approach to interpretation and visualization when dealing with a posterior whose support is a singular, curved manifold is to appeal to Riemannian geometry; a good reference is Pennec (1999, 2006). The idea is to represent the manifold as a flat space⁶

⁶The terms flat space and Euclidean space are used interchangeably.

called a chart. For example, if the manifold were the surface of the earth with elevations disregarded, a chart would be a two dimensional world map and the point cloud would map to points on this world map.

When the chart is a flat space, one can define or infer probability distributions on the chart following standard statistical methods for a Euclidean space and then map them to distributions on the manifold M (Pennec, 2006). Similarly, one can determine credibility regions on the chart and then map them to credibility regions on the manifold.

The flat space one uses as a chart with the Riemannian approach is the plane $T_{\bar{x}}M$ tangent to the manifold M at the intrinsic mean \bar{x} . A point \mathbf{x}_i from $\mathcal{D} \subset M$ is plotted on this chart as follows. One presumes that an analytic representation of a geodesic of the form $\gamma(t)$ is available. For each geodesic $\gamma(t)$ with $\gamma(0) = \bar{x}$, the tangent vector $\frac{d}{dt}\gamma(0)$ is in $T_{\bar{x}}M$. Let $\gamma_i(t)$ be the geodesic connecting \bar{x} to the point \mathbf{x}_i for which the distance $\delta_G(\bar{x}, \mathbf{x}_i)$ is smallest and let \hat{v}_i be the tangent vector $\hat{v}_i = \frac{d}{dt}\gamma_i(0)$. The marker \hat{z}_i corresponding to \mathbf{x}_i is placed on the chart $T_{\bar{x}}M$ at $\hat{z}_i = \delta_G(\bar{x}, \mathbf{x}_i) \frac{\hat{v}_i}{\|\hat{v}_i\|}$.

There are some technical problems with this approach, the most important of which is that the geodesic with smallest distance may not be unique. These can be addressed in ways that need not concern us because the method is not feasible when all one has available is a point cloud on M . But we can borrow the ideas of a chart, of mapping \mathbf{x}_i from the cloud to \hat{z}_i on the chart by means of lines emanating from \bar{x} , and of making the straight line distance of \hat{z}_i from \bar{x} along the line in the chart the same as the geodesic distance $\delta_G(\bar{x}, \mathbf{x}_i)$ from \bar{x} to \mathbf{x}_i on the manifold.

From the present perspective the most serious deficiency of the Riemannian approach is making the chart tangent to the manifold M at the intrinsic mean \bar{x} . The coordinate system on such a chart has basis vectors that are orthogonal to the Jacobian of $q(\bar{x})$. This interferes with interpretation because coordinates on the chart are not expressed in the natural parameters $\mathbf{x} = (\rho, \theta)$. Therefore, the chart that we shall use is a plane of dimension d whose coordinates are a subset of \mathbf{x} . There appears to be no defensible best way to construct a chart with an interpretable coordinate system nor is it clear that uniqueness even matters. The chart suggested here is constructed as follows.

One sets a ball $B \subset \mathbb{R}^{d_a}$ centered at \bar{x} that contains about 5% of the points in \mathcal{D} . Only

the points from \mathcal{D} in B are used to find the basis vectors for the chart so as to eliminate the influence of points on the fringe of the point cloud \mathcal{D} . Recall that the number of restrictions m is known. The goal is to find that set of d explanatory variables that best explain the remaining set of m variables.

The only practical effect of changing the above suggestion of using 5% of the points in \mathcal{D} is to possibly change which set of d explanatory variables get chosen. In this connection, one might override the automatic selection suggested immediately below in order, e.g., to ensure that all location parameters in ρ get selected.

Denote a set of m indices from the set $\mathcal{I} = \{1, 2, \dots, d_a = m + d\}$ by LHS_k and denote the set of those that remain by RHS_k ; thus, $\text{LHS}_k \cup \text{RHS}_k = \mathcal{I}$ and there are $\binom{d_a}{m}$ distinct pairs $(\text{LHS}_k, \text{RHS}_k)$. Regress each variable in LHS_k on the variables in RHS_k with an intercept term. Denote the sum of the R^2 for these m regressions by R_k^2 . Let \hat{k} index the maximum of the R_k^2 . The d elementary vectors e_j with $j \in \text{RHS}_{\hat{k}}$ are the basis vectors for the chart. Denote the d_a by d matrix containing the e_k as columns by $P_{\bar{x}}$ and the chart by $\mathcal{C}_{\bar{x}}$.

A point \mathbf{x}_i from from $\mathcal{D} = \{\mathbf{x}_i\}_{i=1}^N \subset M$ is plotted on the chart $\mathcal{C}_{\bar{x}}$ as follows. Put $v_i = P_{\bar{x}}P_{\bar{x}}^\top(\mathbf{x}_i - \bar{\mathbf{x}})$; v_i is the orthogonal projection of $\mathbf{x}_i - \bar{\mathbf{x}}$ onto $\mathcal{C}_{\bar{x}}$. The marker \mathbf{z}_i corresponding to \mathbf{x}_i is placed on the chart $\mathcal{C}_{\bar{x}}$ at $\mathbf{z}_i = \delta_G(\bar{\mathbf{x}}, \mathbf{x}_i) \frac{v_i}{\|v_i\|}$. The marker $\bar{\mathbf{z}}$ for $\bar{\mathbf{x}}$ is placed at zero.

Note that the coefficients $c \in \mathbb{R}^d$ for $\mathcal{C}_{\bar{x}} = \{z = P_{\bar{x}}c \mid c \in \mathbb{R}^d\}$ show up as the non-zero elements of $z \in \mathcal{C}_{\bar{x}}$.

Denote probability with respect to the empirical distribution on the chart by

$$P_{\mathcal{C}_{\bar{x}}}(R \mid z) = \frac{1}{N} \sum_{i=1}^N I(\mathbf{z}_i \in R). \quad (15)$$

The draws $\mathcal{D} = \{\mathbf{x}_i\}_{i=1}^N$ that are expressed in terms of the natural parameter space are in one-to-one correspondence with the points $\{\mathbf{z}_i\}_{i=1}^N$ on the chart $\mathcal{C}_{\bar{x}}$ that are expressed in terms of the basis vectors $P_{\bar{x}}$. Denote this mapping by

$$\mathbf{Z}(\mathbf{x}_i) \mapsto \mathbf{z}_i \in \mathcal{C}_{\bar{x}} \quad \mathbf{X}(\mathbf{z}_i) \mapsto \mathbf{x}_i \in \mathcal{D} \quad (16)$$

A credibility region constructed on the chart, which is a flat space and therefore amenable to standard statistical analysis, can be mapped to the manifold M in \mathbb{R}^{d_a} by mapping the points \mathbf{z}_i in the credibility region on the chart to the corresponding points \mathbf{x}_i on the manifold

M using (16). As a practical matter, for reporting results when $d_a > 3$, one will probably have to restrict attention to regions of the form (11) that can be described in tabular form.

4 Interpretation

The main contribution of this paper is to study the consequences of departing from the conventions of the related literature by substituting a chart with an interpretable coordinate system for the conventional chart. As noted earlier, the reference system for the point cloud \mathcal{D} is extrinsic; that is, the parameters $\mathbf{x} = (\rho, \theta)$ are expressed in a d_a -dimensional, Euclidean coordinate system that embeds the d -dimensional manifold M . This permits using a subset of the basis vectors for the parameter space as the basis vectors for the chart thereby providing an interpretable coordinate system for the chart.

Due to the one-to-one correspondence (16) between the chart $\mathcal{C}_{\bar{\mathbf{x}}}$ and the manifold M , an interpretable coordinate system on the chart is an aid to visualization: The x and y axes of three dimensional plots become meaningful relative to the extrinsic coordinate system.

In the related literature, the appeal of Riemannian geometry grows out of cartography, the study of distributions on the sphere, and, to some extent, familiarity with general relativity. At least in cartography, being on the surface of the manifold makes sense: A slight error in someone's location puts them in the air falling to death or underground suffocating to death. Slight differences matter.

For the motivating problem, slight differences do not matter. The functions $f(y|x, \rho)$ given by (1) and $\gamma(\rho, \theta)$ given by (2) can accept any values in the parameter space that satisfy support conditions. These values do not have to be on the manifold M given by (5).

5 Curvature and Visualization

5.1 Assessing Curvature

The idea here is to compare the Euclidean distance $\delta_E(\bar{\mathbf{x}}, \mathbf{x}_i)$ between the intrinsic mean $\bar{\mathbf{x}}$ and a point $\mathbf{x}_i \in \mathcal{D}$ to the geodesic distance $\delta_G(\bar{\mathbf{x}}, \mathbf{x}_i)$ between these two points. Curvature is indicated when the geodesic distance exceeds the Euclidean distance. For this to be effective, the points \mathbf{x}_i that are selected for examination must be exhaustive and must be ordered such

that a graphical display is interpretable.

Let

$$Q_i = (\mathbf{x}_i - \bar{\mathbf{x}})^\top V^{-1}(\mathbf{x}_i - \bar{\mathbf{x}}) \quad (17)$$

for $x_i \in \mathcal{D}$ and let

$$\mathcal{Q} = \{Q_i \mid \mathbf{x}_i \in \mathcal{D}\}. \quad (18)$$

Select for examination those \mathbf{x}_i for which Q_i is between the 90th and 95th quantiles of \mathcal{Q} .⁷ Denote the set of such \mathbf{x}_i by \mathcal{L} . Select some \mathbf{x}^* from \mathcal{L} as a benchmark and order the points in \mathcal{L} by $\cos(\mathbf{x}^*, \mathbf{x}_i)$, largest to smallest.⁸ Let y_j , $j = 1, \dots, K$, denote these ordered points. On the same graph plot $\delta_G(y_i, \bar{\mathbf{x}})$ against $i = 1, \dots, K$ and $\delta_E(y_j, \bar{\mathbf{x}})$ against $j = 1, \dots, K$. An example is Panel (a) of Figure 5. The vertical distance between these two lines is a measure of curvature.

5.2 Visualization

There is a large literature on visualization of high dimensional data. A good survey is Engle, Huttenberger, and Hamann (2011). Unfortunately, in general, this literature can be classified as cluster analysis and therefore is not very relevant to our problem.

For $d_a = 3$ visualization is trivially easy; see Panels (b), (c), and (d) of Figures 2 and 4 which plot the manifold M in Panel (b), credibility region R_r in (c), and region R_e in (d). Note that the x-axis and y-axis must be chart variables. The variable on the z-axis is a predicted value from a regression of $\mathbf{x}_i = X(\mathbf{z}_i)$ on $\mathbf{z}_i \in \mathcal{C}_{\bar{\mathbf{x}}}$; see the figure legends.

The idea is to borrow from these plots, put meaningful groups of chart variables on the x and y axes, and a relevant predicted manifold variable on the z axis, which is usually an element of θ . Figure 5, which relates to Example 7.3, is a good illustration of this idea. In Panels (b), (c), and (d), the x-axis is location, the y-axis is heteroskedasticity, the scale chart variables are held fixed, and the z-axis is price elasticity.

⁷Other choices such as the 50th and 52th quantiles work equally well.

⁸ $\cos(u, v) = u^\top v / (\|u\| \|v\|)$.

6 Credibility Intervals

Because a joint credibility rectangle respects the fact that the point cloud is confined to the manifold M , we suggest using joint credibility rectangles rather than marginal credibility intervals. We construct the chart rectangle as follows:

On the chart, compute τ such that

$$P_{\mathcal{C}_{\bar{x}}}(\tilde{R}_\tau | z) = 1 - \alpha$$

where

$$\tilde{R}_\tau = \times_{k=1}^d [\bar{z}_k - \tau \text{sdev}(\mathbf{z}_{kk}), \bar{z}_k + \tau \text{sdev}(\mathbf{z}_{kk})]. \quad (19)$$

In (19), k denotes an element of \bar{z} , \bar{z} is the point on the chart to which \bar{x} maps, and the $\text{sdev}(\mathbf{z}_{kk})$ are the square roots of the diagonal elements of

$$\tilde{V} = \frac{1}{N} \sum_{i=1}^N (\mathbf{z}_i - \bar{z})(\mathbf{z}_i - \bar{z})^\top \quad (20)$$

for $\mathbf{z}_i \in \mathcal{C}_{\bar{x}}$.

For tabular display of the chart credibility rectangle, we shall have to shift the center, viz.

$$\hat{R}_\tau = \times_{k=1}^d [\bar{x}_k - \tau \text{sdev}(\mathbf{z}_{kk}), \bar{x}_k + \tau \text{sdev}(\mathbf{z}_{kk})], \quad (21)$$

where \bar{x} is the intrinsic mean. This is the formula used in columns labeled ‘‘Chart’’ in Tables 1, 2, and 4.

The corresponding credibility rectangle on the parameter space is

$$R_\tau = \{\mathbf{x}_i | \mathbf{x}_i = X(\mathbf{z}_i), \mathbf{z}_i \in \tilde{R}_\tau\}.$$

Unfortunately, it is hard to find the boundaries of R_τ using this construction. One can estimate the boundaries of R_τ by $\min\{\mathbf{x}_{ki}\}$ and $\max\{\mathbf{x}_{ki}\}$ computed element by element, $k = 1, \dots, d_a$, over $\mathbf{x}_i \in R_\tau$ but usually these estimates are noisy. Alternatively one can use the regression of $\mathbf{x}_i = X(\mathbf{z}_i)$ on $\mathbf{z}_i \in \mathcal{C}_{\bar{x}}$ discussed in Subsection 5.1 to map the boundary of \tilde{R}_τ to the parameter space but this suffers from prediction error and specification error. These boundary errors are not of much concern for graphical display but are problematic for tabular reporting.

For tabular display one is better advised to compute R_τ using (11), (10), and (7). But, as seen later, these intervals will be shorter than those constructed according to the previous paragraph.

7 Examples

7.1 A Flat Manifold Example

Figure 1 displays an iid point cloud on a 1D flat manifold embedded in a 3D Euclidean parameter space. The manifold is a one dimensional circle tilted with respect to the parameter space. Figure 2 displays the curvature indicators discussed in Subsections 5.1 and 5.2. The geodesic and Euclidean lines in Panel (a) have negligible separation which implies that the manifold is a flat space. Panel (b) displays the manifold variable labeled `var_z` as a regression prediction from chart variables that are labeled `var_x` and `var_y`. The enumeration on the right hand side of `var_x` and `var_y` on the axes of Figure 1 refers to the coordinates of the parameter space, not to coordinates of the chart, as does, perforce, the enumeration for `var_z`. Panel (c) shows the intersection of the 95% rectangular credibility region R_r given by (11) with the regression surface shown in Panel (b). Panel (d) is the same for R_e given by (12).

Figure 1 about here.

Figure 2 about here.

7.2 A Curved Manifold Example

Figure 3 displays an iid point cloud on a 1D curved manifold embedded in a 3D Euclidean parameter space. The manifold is a one dimensional sphere. Figure 4 displays the curvature indicators discussed in Subsections 5.1 and 5.2. The geodesic and Euclidean lines in Panel (a) have modest separation which implies that the manifold is a mildly curved space. Panel (b) displays the manifold as a regression prediction from chart variables, labeled `var_x` and `var_y`, to a manifold variable, `var_z`. Their enumeration refers to coordinates of the natural parameter space as described above. Panel (c) shows the intersection of the 95% rectangular

credibility region R_r given by (11) with the regression surface shown in Panel (b). Panel (d) is the same for R_e given by (12).

Figure 3 about here.

Figure 4 about here.

Table 1 about here.

7.3 A Simple Demand and Supply Example

This example is from Gallant (2022a, 2022b) where several measures of location and scale are reported that make useful comparison with results reported here. It is also an instance of Example 3 of Bornn, Shephard, and Solgi, (2018).

The demand and supply system is

$$x_t = (\sigma_x + \rho_x x_{t-1}) z_{1,t} \quad (22)$$

$$\log q_{d,t} = a_1 + a_2 \log p_t + \sigma_d z_{2,t} \quad (23)$$

$$\log q_{s,t} = b_1 + b_2 \log p_t + x_t + \sigma_s z_{3,t} \quad (24)$$

with solution $(\log p_t, \log q_t)$ under $q_t = q_{d,t} = q_{s,t}$, where $\sigma_x = 3$, $\rho_x = 0.2$, $a_1 = 12$, $a_2 = -2$, $b_1 = 3$, $b_2 = 4$, $\sigma_d = \sigma_s = 0.1$, $z_{i,t}$ standard normal, and sample size $n = 500$. Note that the supply sifter x_t is heteroscedastic with variance dependent on x_{t-1} whence the same for price p_t and quantity q_t . The data are $y_t = (\log p_t, \log q_t, x_t)$ for $t = 1, 2, \dots, n$.

The likelihood used for estimation is normal with heteroscedastic errors that depend on past values of y_t :

$$y_t \sim n_3(y_t | \mu, \Sigma_{t-1}) \quad (25)$$

$$\Sigma_{t-1} = RR^\top + P(y_{t-1} - \mu)(y_{t-1} - \mu)^\top P^\top, \quad (26)$$

where R is upper triangular, and P is diagonal. Thus,

$$\rho = (\mu_1, \mu_2, \mu_3, R_{1,1}, R_{1,2}, R_{2,2}, R_{1,3}, R_{2,3}, R_{3,3}, P_{1,1}, P_{2,2}, P_{3,3}) \in \mathbb{R}^{12}.$$

A set of moment conditions for estimation of the demand equation (23) are

$$m_{d,1}(y_t, y_{t-1}, \rho, \theta) = \log q_t - a_1 - a_2 \log p_t \quad (27)$$

$$m_{d,2}(y_t, y_{t-1}, \rho, \theta) = x_t m_{d,1}(y_t, y_{t-1}, \rho, \theta) \quad (28)$$

$$m_{d,3}(y_t, y_{t-1}, \rho, \theta) = x_{t-1} m_{d,1}(y_t, y_{t-1}, \rho, \theta) \quad (29)$$

$$\theta = (a_1, a_2)$$

$$\rho \quad \text{not used}$$

The prior for ρ is independent normal with location the unconstrained maximum likelihood estimates of (25) and scale twice the maximum likelihood standard deviation. The prior for $\theta = (a_1, a_2)$ is independent normal with means $(12, -2)$ and standard deviations $(2, 2)$. The support conditions on R and P of (26) are that diagonals of R must be positive, the first diagonal element P must be positive, and the eigenvalues of the companion matrix of Σ_{t-1} must be less than one in absolute value. In addition, a_1 must be positive and a_2 negative. The point cloud is 50,000 draws generated using Gallant (2020) with tuning parameters as in Gallant (2022a).

Curvature is moderate, as seen in panel (a) of Figure 5, which plots the explicit the curvature measure proposed in Section 5.1. Regression visualizations are shown in the remaining panels, which plot price elasticity a_2 , z-axis, confined to the manifold M against location chart variables moved as a group, x-axis, and heteroskedasticity chart variables moved as a group, x-axis, Scale parameters are held fixed at \bar{z} .

Figure 5 about here.

Table 2 about here.

7.4 Hansen and Singleton (1982) Revisited

The original data from Hansen and Singleton (1982) are published in Gallant (1987, pp. 416–420) and available online at www.aronaldg.org/webfiles/data/hansena.dat. The data are monthly observations, $t = 1, \dots, 239 = n$, for the US economy on nondurables and

services, NDS_t , population, POP_t , value weighed NYSE returns with dividends, $NYSE_t$, and the implicit price deflator, DFL_t , from 1959 through 1978. Set

$$\begin{aligned} \text{lcg} &= \log \left[\frac{(NDS_t)/(POP_t)}{(NDS_{t-1})/(POP_{t-1})} \right] \\ \text{lsr} &= \log \left[(1 + NYSE_t) \frac{(DFL_{t-1})}{(DFL_t)} \right]. \end{aligned}$$

The economic model describes asset pricing in an exchange economy under constant relative risk aversion, CRRA, utility.

The statistical model is equations (1) through (4) where $f(y_t | x_{t-1}, \rho)$ is a bivariate, Gaussian, conditionally heteroskedastic, autoregression with location and scale given by

$$\mu_t = b + By_{t-1} \quad (30)$$

$$\Sigma_t = RR^\top + Q\Sigma_{t-1}Q^\top + P(y_{t-1} - b - By_{t-2})(y_{t-1} - b - By_{t-2})^\top P^\top \quad (31)$$

where R is upper triangular, P is diagonal, and Q is a scalar. Thus

$$\rho = (b_1, b_2, B_{1,1}, B_{2,1}, B_{1,2}, B_{2,2}, R_{1,1}, R_{1,2}, R_{2,2}, P_{1,1}, P_{2,2}, Q_{1,1}) \in \mathbb{R}^{12}. \quad (32)$$

Set

$$y_t = \begin{pmatrix} y_{1t} \\ y_{2t} \end{pmatrix} = \begin{pmatrix} \text{lsr}_t \\ \text{lcg}_t \end{pmatrix}. \quad (33)$$

The variable x_{t-1} denotes the information set and has as many lags in it as are needed to compute $f(y_t | x_{t-1}, \rho)$ of equation (1).

The moment conditions for estimating $\theta = (\beta, \gamma)$ are

$$\mathbf{m}(y_t, x_{t-1}, \theta) = \begin{pmatrix} 1 \\ \text{lsr}_{t-1} \\ \text{lcg}_{t-1} \end{pmatrix} [1 - \exp(\log \beta - \gamma \text{lcg}_t + \text{lsr}_t)], \quad (34)$$

for $t = T_0, \dots, n$. Thus, (2) becomes

$$0 = q(\rho, \theta) = \frac{1}{n} \sum_{t=T_0}^n \int \mathbf{m}(y, x_{t-1}, \theta) f(y | x_{t-1}, \rho) dy, \quad q \in \mathbb{R}^m \quad (35)$$

where ρ is (32) above and $\theta = (\beta, \gamma)$.

The prior is shown in Table 3. The support conditions are that $0.1 < \beta < 1.0$; $0 < \gamma < 200$; $R_{1,1}$ and $R_{2,2}$ must be positive; $P_{1,1}$ and $Q_{1,1}$ must be positive; and the eigenvalues of the companion matrices for the location function and the scale function must be less than one in absolute value.

Parameter estimates of θ from Hansen and Singleton (1982), unconstrained maximum likelihood estimates of ρ using Gallant and Tauchen (1990), and Bayesian estimates of (ρ, θ) subject to (2), (3), and (4) using Gallant (2020), together with confidence intervals for the frequentist estimators and credibility intervals for the Bayes estimator are shown in Table 4. The Bayes estimates differ little from the Hansen and Singleton estimates.

Table 5 shows marginal credibility intervals for comparison with the joint credibility intervals shown in Table 4. As shown in the table, the length penalty is about 50% for using joint intervals instead of marginal intervals. The view advanced here is that joint intervals are preferred because they reflect the fact that the support of the posterior is a curved, singular manifold and because their extra length partially compensates for ignoring geodesic distance in their construction.

The likelihood ratio test of the restriction (35) is $\lambda = 1319.57 - 1289.92 = 29.65$, which is asymptotically chi square on one degree of freedom provided one agrees that the prior, Table 3, is loose relative to the credibility intervals shown in Table 4. One rejects the restrictions (35) at the $\alpha = 0.005$ significance level.

Curvature is strong, as seen in panel (a) of Figure 6, which plots the explicit the curvature measure proposed in Section 5.1. Regression visualizations are shown in the remaining panels, which plot γ , the z-axis, confined to the manifold M against chart location variables moved as a group, the y-axis, and chart scale variables moved as a group, the x-axis,

Table 3 about here.

Table 4 about here.

Table 4 about here.

Figure 6 about here.

8 Conclusion

The notion of a chart is a standard adjunct to the interpretation of a point cloud confined to a curved, singular manifold that is embedded in a Euclidean space. Here, the points are drawn from a posterior and the embedding space is the parameter space. A chart is a Euclidean space of the same dimension as the manifold. Points map one-to-one between the manifold and the chart. Euclidean distance on the chart equals geodesic distance on the manifold. In Riemannian geometry the chart is tangent to the manifold with the consequence that chart coordinates are not easily interpretable with respect to the parameter space. This paper proposes a chart that uses a subset of parameter space basis vectors but is otherwise similar to the Riemannian chart. This is a substantial aid to visualization by means of interpretable 3D plots.

Using a discrete analog of a space filling curve, a measure of curvature is proposed. The curve sweeps through points in an elliptical band at the fringe of the point cloud. Curvature is measured by the difference between geodesic distance from a point in the band to the intrinsic mean and Euclidean distance for same.

As a preferred reporting mechanism, a system of simultaneously valid credibility rectangles is proposed. It can be constructed on the chart then mapped to the manifold or constructed directly on the manifold. Comparing the two constructions provides a secondary measure of curvature.

The proposed methods are illustrated by examples that include replication of the classic Hansen and Singleton (1982) paper using their original data. What is clear from the examples is that confidence rectangles that use geodesic distance are improbably large. Confidence rectangles constructed on the point cloud using Euclidean distance appear reasonable relative to other statistical methods.

9 Supplementary Materials

C++ and R code: Provided is the code needed to produce Table 2 and Figure 5 from the included point cloud. (`curve_ds_eg.zip`, a MAC zipped directory, or from www.aronaldg.org/webfiles/curve_ds_eg)

10 References

- Bornn, Luke, Neil Shephard, and Reza Solgi (2018), “Moment Conditions and Bayesian Nonparametrics,” *Journal of the Royal Statistical Society, Series B* 81(1), 5–43.
- Dijkstra, E. W. (1959), “A Note on Two Problems in Connection with Graphs,” *Numerische Mathematik* 1 269–271
- Engel, Daniel, Lars Huttenberger, and Bernd Haman (2011); “A Survey of Dimension Reduction Methods for High-dimensional Data Analysis and Visualization,” in *Proceedings of IRTG 1131, “Visualization of Large and Unstructured Data Sets Workshop 2011,”* Christoph Garth, Ariane Middel, and Hans Hagen, Eds.; pp. 135–149 <https://drops.dagstuhl.de/opus/volltexte/2012/3747>
- Gallant, A. Ronald (1987), *Nonlinear Statistical Models*, John Wiley and Sons, New York,
- Gallant, A. Ronald, (2020), “NPB: Nonparametric Bayes Subject to Overidentified Moment Conditions, Version 1.3, User’s Guide,” Manuscript, www.aronaldg.org/webfiles/npb/
- Gallant, A. Ronald (2022a), “Nonparametric Bayes Subject to Overidentified Moment Conditions,” *Journal of Econometrics* 228, 27–38.
- Gallant, A. Ronald (2022b), “Variance-Covariance from a Metropolis Chain on a Curved, Singular Manifold,” Forthcoming, *Journal of Econometrics*.
- Gallant, A. Ronald, and Douglas W. Nychka (1987), “Semi-Nonparametric Maximum Likelihood Estimation,” *Econometrica* 55, 363–390.
- Gallant, A. Ronald, and George Tauchen (1989), “Seminonparametric Estimation of Conditionally Constrained Heterogeneous Processes: Asset Pricing Applications,” *Econometrica* 57, 1091–1120.
- Gallant, A. Ronald, and George Tauchen (1990), “SNP: A Program for Nonparametric Time Series Analysis, Version 9.1, User’s Guide,” , Manuscript, www.aronaldg.org/webfiles/snp/.

- Hansen, Lars P., and Kenneth J. Singleton (1982), “Generalized Instrumental Variable Estimation Estimators of Nonlinear Rational Expectations Models,” *Econometrica* 50, 1269–1286.
- Memoli, Facundo, and Guillermo Sapiro (2001), “Fast Computation of Weighted Distance Functions and Geodesics on Implicit Hyper-surfaces,” *Journal of Computational Physics* 173, 730–764.
- Morgan, Frank (2016), *Geometric Measure Theory, A Beginner’s Guide, 5th Edition*, San Diego CA, Academic Press.
- Penec, Xavier (1999), “Probabilities and Statistics on Riemannian Manifolds: Basic Tools for Geometric Measurements. In A.E. Cetin, L. Akarun, A. Ertuzun, M.N. Gurcan, and Y. Yardimci, editors, *Proceedings of Nonlinear Signal and Image Processing 1*, 194–198, June 20-23, Antalya, Turkey, 1999. IEEE-EURASIP. Penec, Xavier (2006), “Intrinsic Statistics on Riemannian Manifolds: Basic Tools for Geometric Measurements,” *Journal of Mathematical Imaging and Vision* 25, 127-154.
- Sethian, J. A. (1996), “A Fast Marching Level Set Method for Monotonically Advancing Fronts,” *Proceedings of the National Academy of Sciences* 93, 1591–1595.
<https://doi.org/10.1073/pnas.93.4.1591>
- Shin, Minchuyul (2015), “Bayesian GMM,” Working paper, Department of Economics, University of Illinois. http://www.econ.uiuc.edu/~mincshin/BGMM_ver05
- Wikipedia (2021), “Frechet Mean,” Last modified June 3, 2021.
https://en.wikipedia.org/wiki/Frechet_mean.
- Zappa, Emilio, Miranda Holmes-Cerfon, and Jonathan Goodman (2018), “Monte Carlo on Manifolds: Sampling Densities and Integrating Functions,” *Communications on Pure and Applied Mathematics* 71, 2609–2647.

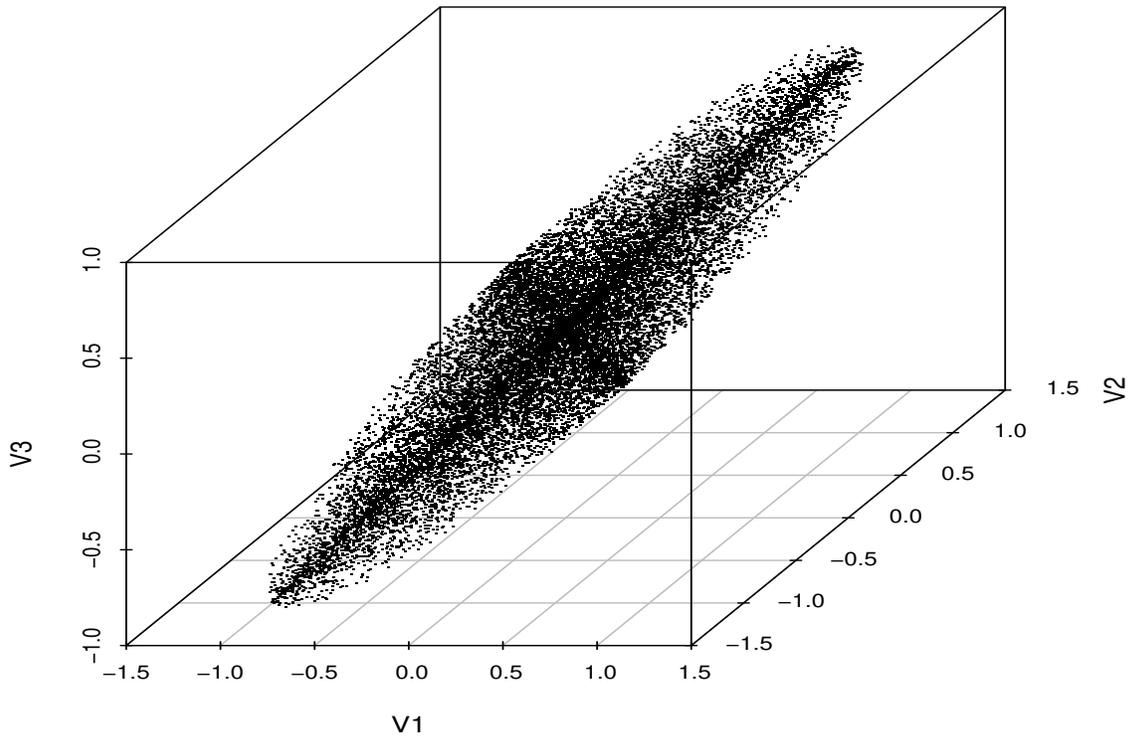


Figure 1. A Point Cloud on a Flat Manifold. An independent and identically distributed sample supported on a one dimensional circle that is tilted with respect to the axes of the three dimensional embedding space. The data are 20,000 iid $uv, v \sim N_3(0, I), u \sim U(0, 1)$ normalized to have length one projected onto two dimensions.

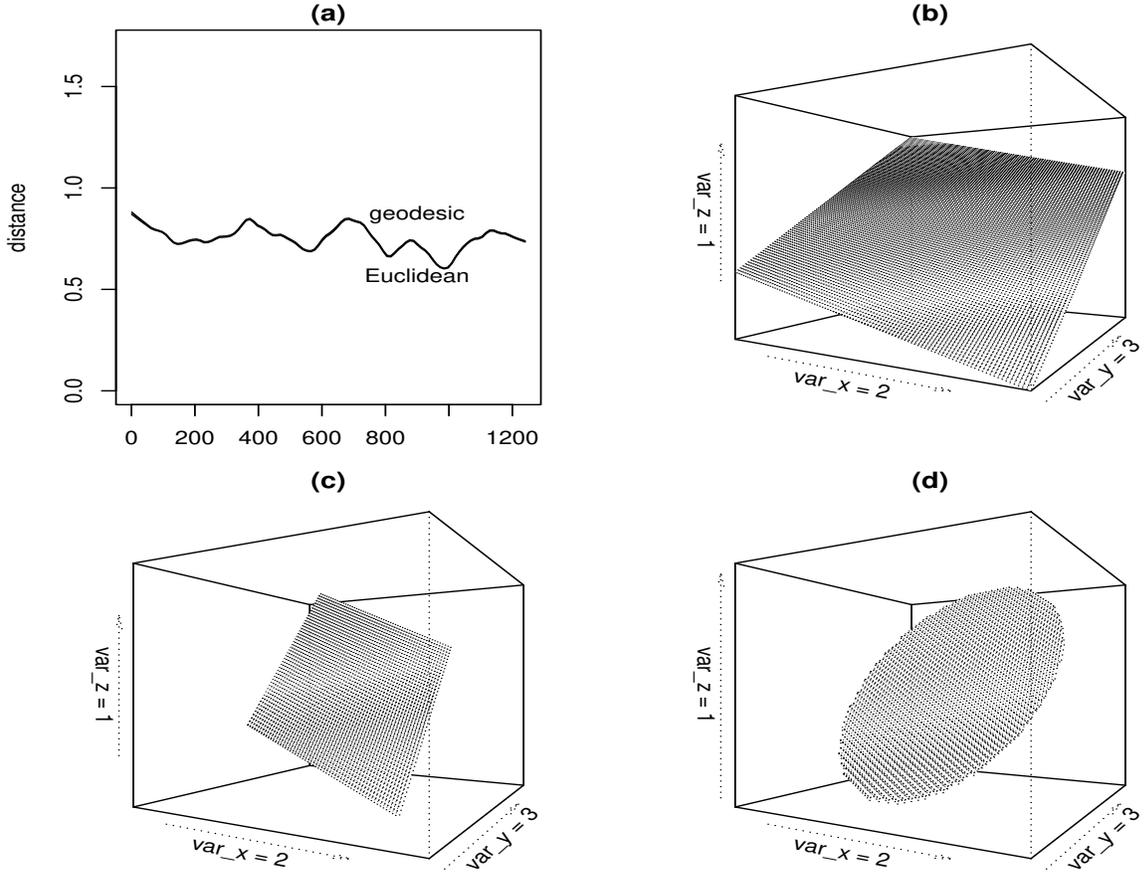


Figure 2. Curvature Indicators for the Flat Manifold Example.

Panel (a). Geodesic and Euclidean Distance. For points $\mathbf{x}_i \in \mathcal{D}$ between the 90% and 95% credibility region boundaries one of them, \mathbf{x}^* , is selected as a benchmark. The points are ordered from largest to smallest according to $\cos(\mathbf{x}^*, \mathbf{x}_i)$. Let $\{\mathbf{y}_j\}_{j=1}^K$ denote the ordered points and note that $\mathbf{y}_1 = \mathbf{x}^*$. Plotted are $\delta_G(\mathbf{y}_j, \bar{\mathbf{x}})$ against $j = 1, \dots, K$, labeled geodesic, and $\delta_E(\mathbf{y}_j, \bar{\mathbf{x}})$, labeled Euclidean; i.e. the horizontal axis is $j = 1, \dots, K$ and the vertical axis is δ_G and δ_E . The distance between the two plots is a measure of curvature. In this instance the distance is negligible because the manifold has no curvature.

Panel (b). Regression Surface. There is a one-to-one mapping between points on the chart to points on the manifold. Shown is a plot of a linear regression of the manifold points onto the chart points. The corners of the plot extend beyond the boundary of the point cloud. The axes labeled `var_x` and `var_y` are chart variables. The axis labeled `var_z` is an embedding variable. Their enumeration refers to coordinates of the natural parameter space.

Panel (c). Rectangular Credibility Region. Shown is the intersection of the 95% rectangular credibility region R_r given by (11) with the regression surface shown in panel (b).

Panel (d). Elliptical Credibility Region. Shown is the intersection of the 95% elliptical credibility region R_e given by (12) with the regression surface shown in panel (b).

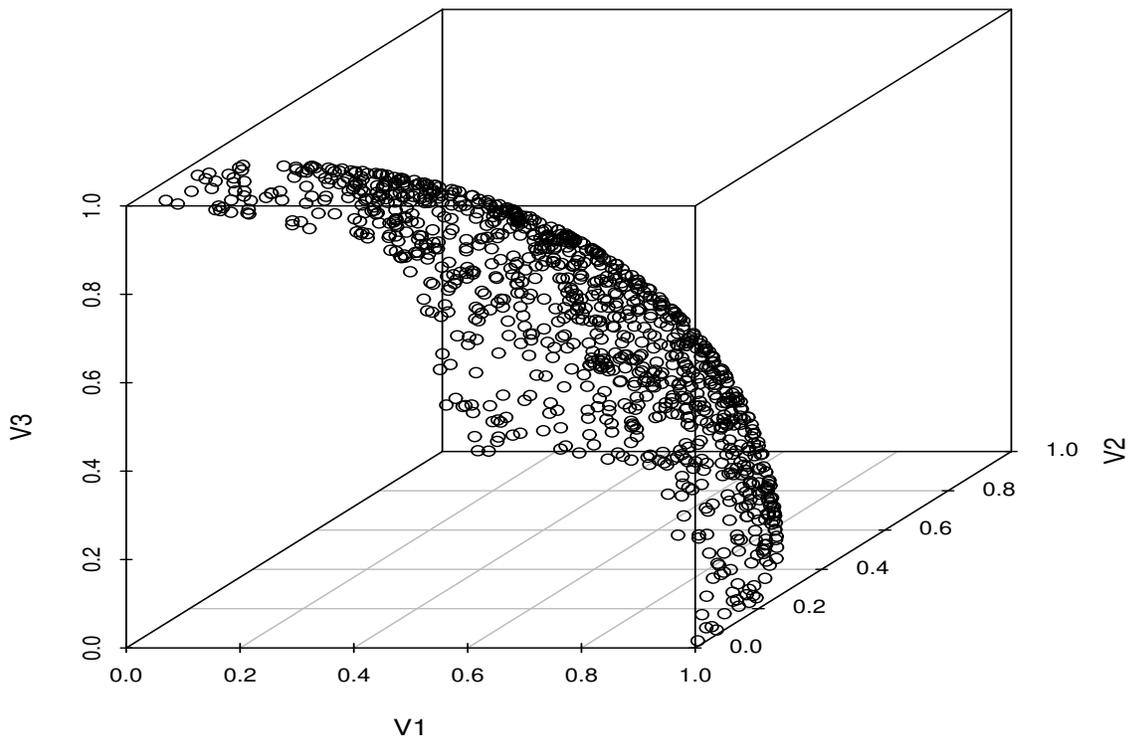


Figure 3. A Point Cloud on a Curved Manifold. An independent and identically distributed sample supported on a one dimensional sphere embedded in a three dimensional parameter space. The data are 20,000 iid $(u_1, u_2, u_3), u_i \sim U(0, 1)$ normalized to have length one.

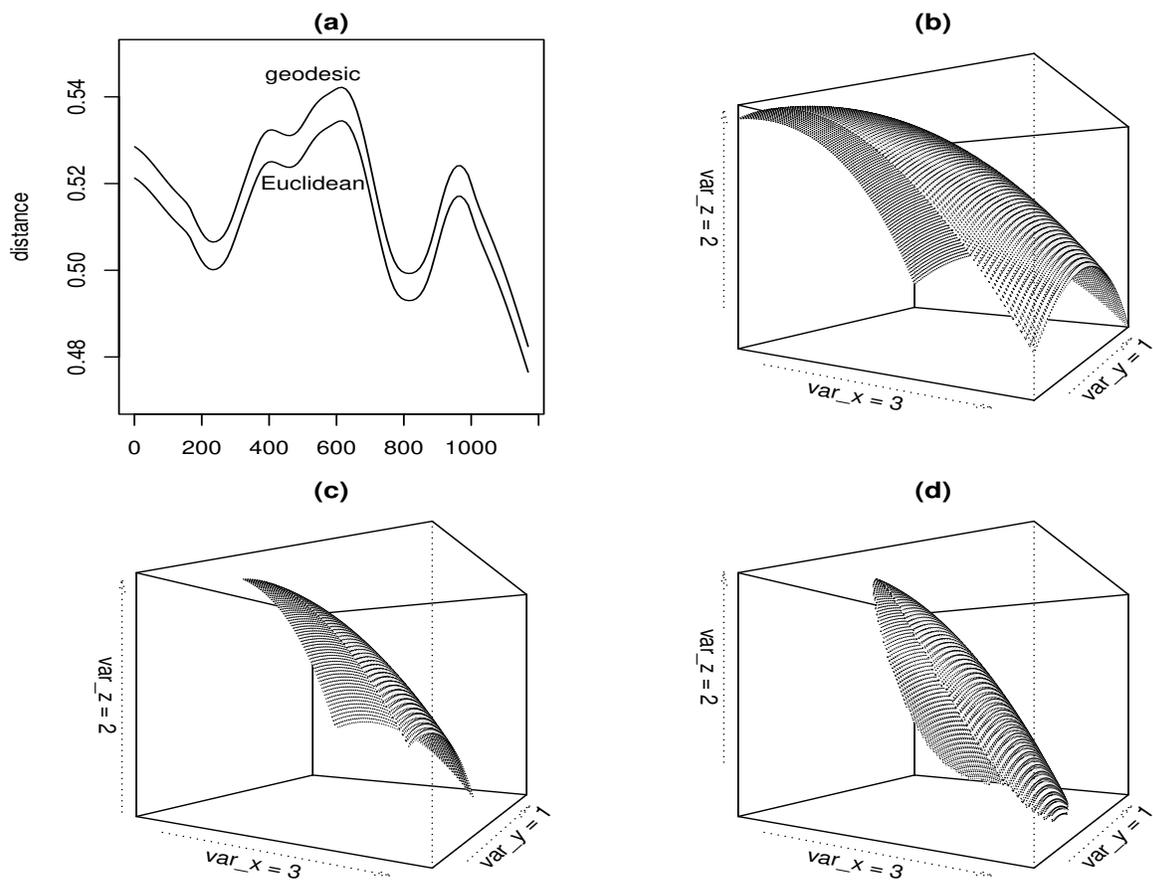


Figure 4. Curvature Indicators for the Curved Manifold Example.

Panel (a). Geodesic and Euclidean Distance. As for Figure 2 except that the two plots are separated by a modest distance that indicates mild curvature.

Panel (b). Regression Surface. As for Figure 2 except that the regression is a quadratic.

Panel (c). Rectangular Credibility Region. As for Figure 2.

Panel (d). Elliptical Credibility Region. As for Figure 2.

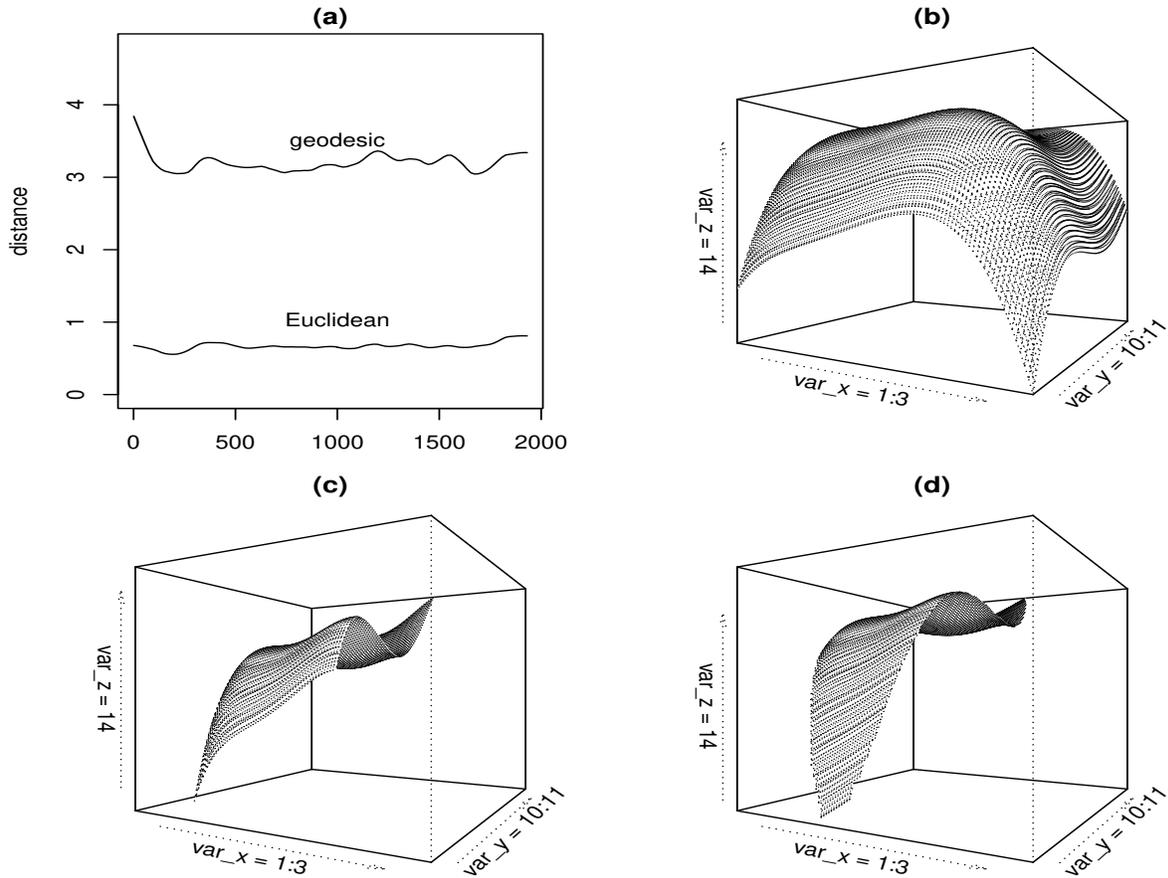


Figure 5. Curvature Indicators for the Demand and Supply Example.

Panel (a). Geodesic and Euclidean Distance. As for Figure 2 except that the two plots are separated by a distance that indicates large curvature.

Panel (b). Regression Surface. As for Figure 2 except that the regression is a degree four polynomial with interactions to degree three. The x-axis is location of which all are on the chart, $\text{var}_x = (\mu_1, \mu_2, \mu_3)$, moved as a group. The y-axis is heteroskedasticity that is on the chart, $\text{var}_y = (P_{1,1}, P_{2,2})$, moved as a group. The z-axis is a_2 , which is price elasticity and is not a chart variable. The scale variables $R_{1,1}$ through $R_{3,3}$ are chart variables that are held fixed at the the intrinsic mean \bar{x} , which plots as zero on the chart.

Panel (c). Rectangular Credibility Region. As for Figure 2.

Panel (d). Elliptical Credibility Region. As for Figure 2.

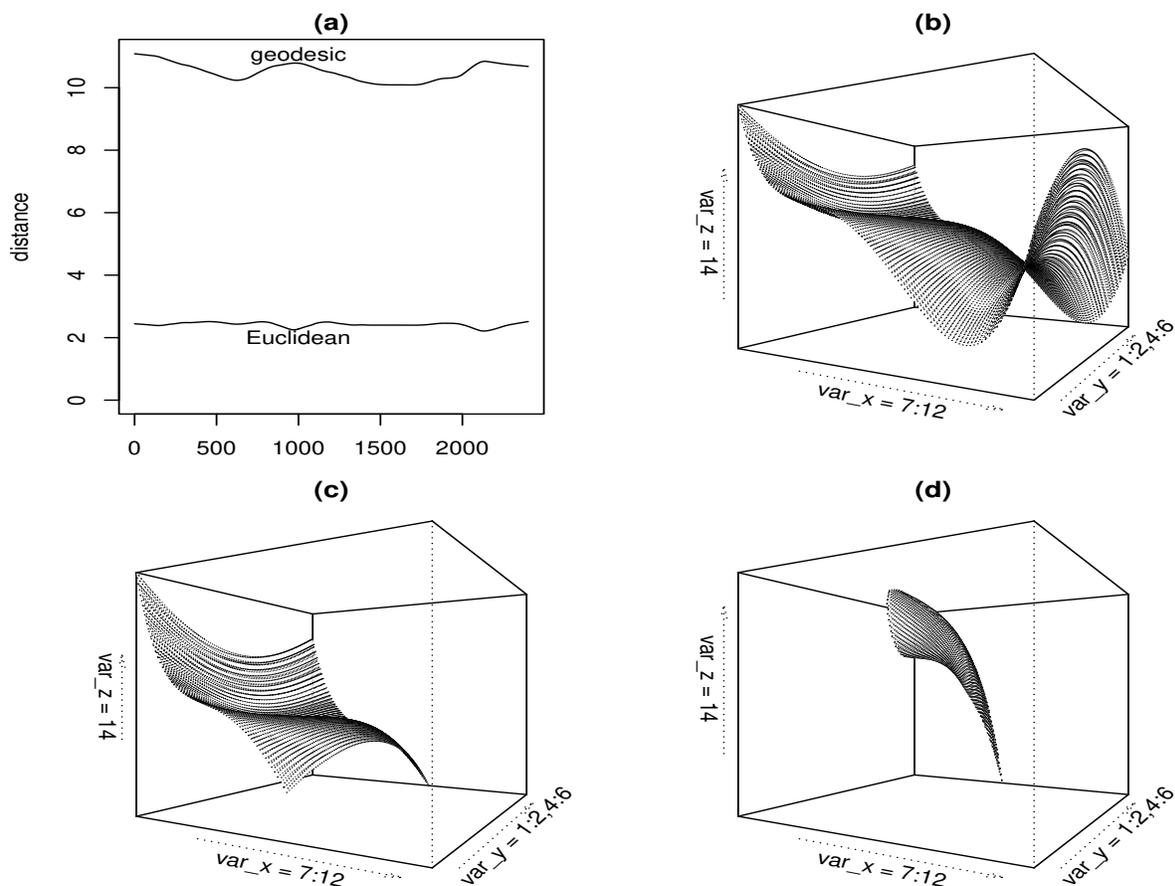


Figure 6. Curvature Indicators for the Hansen and Singleton (1982) Example.

Panel (a). Geodesic and Euclidean Distance. As for Figure 2 except that the two plots are separated by a distance that indicates large curvature.

Panel (b). Regression Surface. As for Figure 2 except that the regression is a degree four polynomial with interactions to degree three. The x-axis is scale that is on the chart, $\text{var}_x = (R_{1,1}, R_{1,2}, R_{2,2}, P_{1,1}, P_{2,2}, Q_{1,1})$, moved as a group. The y-axis is location that is on the chart, $\text{var}_y = (b_1, b_2, B_{2,1}, B_{1,2}, B_{2,2})$, moved as a group. The z-axis is γ , which is not a chart variable.

Panel (c). Rectangular Credibility Region. As for Figure 2.

Panel (d). Elliptical Credibility Region. As for Figure 2.

Table 1. Estimates for the Curved Manifold Example

Parameter	Estimate	Credibility Intervals	
		Cloud	Chart
x	0.57675	0.01608, 1.13742	-0.06157, 1.21507
y	0.57883	0.01830, 1.13937	
z	0.57647	0.01994, 1.13299	-0.05858, 1.21152

The parameter estimate is the intrinsic mean computed using (14). Credibility intervals at 95% were computed using (11), (10), and (7) for the column labeled Cloud and (21), (20), and (15) for the column labeled Chart.

Table 2. Estimates for the Demand and Supply Example

Parm	Max Likelihood		Bayesian		
	Estimate	Conf Interval	Estimate	Cred Intervals	
				Cloud	Chart
μ_1	0.00527	-0.08473, 0.09527	-0.00003	-0.09392, 0.09386	-1.71786, 1.71781
μ_2	-0.00363	-0.09265, 0.08539	-0.00079	-0.10119, 0.09962	-1.58011, 1.57854
μ_3	0.00074	-0.08773, 0.08921	0.01134	-0.09589, 0.11857	-1.59993, 1.62260
$R_{1,1}$	0.98124	0.90443, 1.05805	0.97910	0.88043, 1.07778	
$R_{1,2}$	-0.00117	-0.04797, 0.04563	-0.00555	-0.06276, 0.05165	-0.86920, 0.85809
$R_{2,2}$	0.99011	0.91490, 1.06532	0.96984	0.84134, 1.09835	-1.10177, 3.04146
$R_{1,3}$	-0.00263	-0.04516, 0.03990	-0.00934	-0.06515, 0.04646	-0.88379, 0.86510
$R_{2,3}$	-0.00300	-0.05129, 0.04529	-0.00224	-0.05528, 0.05081	-0.85203, 0.84756
$R_{3,3}$	0.98842	0.92027, 1.05657	1.00550	0.91594, 1.09506	-0.31672, 2.32771
$P_{1,1}$	0.18437	-0.07006, 0.43880	0.16312	-0.07847, 0.40470	-3.77625, 4.10248
$P_{2,2}$	0.15678	-0.12264, 0.43620	0.03334	-0.28706, 0.35375	-3.85954, 3.92623
$P_{3,3}$	-0.14894	-0.38551, 0.08763	-0.00741	-0.39943, 0.38461	-5.19730, 5.18248
a_1			11.98399	11.95268, 12.01531	
a_2			-1.99489	-2.01493, -1.97486	

Maximum likelihood estimates are unconstrained. Confidence intervals are the estimate plus and minus 1.96 the standard error. Bayesian MCMC draws were computed using Gallant (2020). The Bayesian parameter estimate is the intrinsic mean computed using (14). Credibility intervals at 95% were computed using (11), (10), and (7) for the column labeled Cloud and (21), (20), and (15) for the column labeled Chart.

Table 3. Prior for the Hansen and Singleton Example.

Parameter	mean	std. dev.
b_1	0.03933	1.00000
b_2	-0.06346	1.00000
$B_{1,1}$	0.00988	1.00000
$B_{2,1}$	-0.10529	1.00000
$B_{1,2}$	0.01288	1.00000
$B_{2,2}$	-0.18929	1.00000
$R_{1,1}$	0.24539	1.00000
$R_{1,2}$	0.00626	1.00000
$R_{2,2}$	0.24970	1.00000
$P_{1,1}$	0.30396	1.00000
$P_{2,2}$	0.22018	1.00000
$Q_{1,1}$	0.93808	1.00000
β	0.99786	0.01000
γ	1.03705	10.00000

The values shown for the mean are those in the λ -prior method's MCMC chain that were closest to the manifold. They were also used as start values for the NPB iterations. See Gallant (2022a, Section 2.2).

Table 4. Estimates for the Hansen and Singleton Example

Parm	Hansen and Singleton		Max Likelihood		Bayesian		
	Estimate	Conf Interval	Estimate	Conf Interval	Estimate	Cred Intervals	
						Cloud	Chart
b_1			0.02294	-0.108, 0.155	0.08931	-0.045, 0.223	-6.905, 7.084
b_2			-0.00924	-0.137, 0.118	-0.00308	-0.129, 0.123	-6.316, 6.310
$B_{1,1}$			0.08608	-0.069, 0.241	0.00685	-0.001, 0.015	
$B_{2,1}$			-0.05166	-0.179, 0.076	-0.05009	-0.122, 0.021	-3.488, 3.388
$B_{1,2}$			-0.00005	-0.122, 0.122	0.03808	0.021, 0.055	-0.931, 1.007
$B_{2,2}$			-0.25700	-0.389,-0.125	-0.29676	-0.410, -0.184	-6.944, 6.350
$R_{1,1}$			0.31081	0.131, 0.491	0.28310	0.162, 0.404	-5.835, 6.401
$R_{1,2}$			0.00186	-0.036, 0.040	0.01012	-0.023, 0.043	-1.636, 1.656
$R_{2,2}$			0.45081	0.273, 0.629	0.30858	0.185, 0.432	-6.097, 6.714
$P_{1,1}$			0.37613	0.208, 0.544	0.29878	0.148, 0.450	-6.675, 7.273
$P_{2,2}$			0.06396	-0.148, 0.275	0.21700	0.104, 0.330	-5.459, 5.893
$Q_{1,1}$			0.87799	0.781, 0.975	0.91154	0.872, 0.951	-1.236, 3.059
β	0.9979	0.993, 1.00			0.99562	0.989, 1.002	
γ	0.9001	0.287, 1.51			0.97476	0.792, 1.158	

Hansen and Singleton estimates are from Hansen and Singleton (1982, Table I, Line 5). Maximum likelihood estimates are unconstrained. Confidence intervals are the estimate plus and minus 1.96 the standard error. Bayesian MCMC draws were computed using Gallant (2020). The Bayesian parameter estimate is the intrinsic mean computed using (14). Credibility intervals at 95% were computed using (11), (10), and (7) for the column labeled Cloud and (21), (20), and (15) for the column labeled Chart.

Table 5. Marginal Credibility Intervals for the Hansen and Singleton Example.

Parameter	Cred Interval	Rel Length
b_1	0.00475, 0.17388	1.58
b_2	-0.08689, 0.08073	1.51
$B_{1,1}$	0.00133, 0.01236	1.45
$B_{2,1}$	-0.10059, 0.00041	1.42
$B_{1,2}$	0.02304, 0.05312	1.12
$B_{2,2}$	-0.40017,-0.19336	1.10
$R_{1,1}$	0.19965, 0.36654	1.45
$R_{1,2}$	-0.01561, 0.03585	1.29
$R_{2,2}$	0.22180, 0.39537	1.41
$P_{1,1}$	0.20267, 0.39490	1.57
$P_{2,2}$	0.11731, 0.31669	1.13
$Q_{1,1}$	0.88369, 0.93939	1.42
β	0.99181, 0.99942	1.65
γ	0.85985, 1.08968	1.59

Marginal credibility intervals are $I_k = [\bar{x}_k - \tau_k \sqrt{v_{kk}}, \bar{x}_k + \tau_k \sqrt{v_{kk}}]$, where k denotes an element of the intrinsic mean \bar{x} , and the v_{kk} are the diagonal elements of V given by (10). The critical value τ_k is chosen such that $P_{\mathcal{D}}(I_k | x, y) = 0.95$. Shown in the Relative Length column is the ratio of the critical value for the joint credibility region shown in Table 4, 2.5260, divided by τ_k .