uc3m | **Universidad Carlos III de Madrid**

@-Arch*i*vo

This document is published at:

Maxentropic Solutions to a Convex Interpolation Problem Motivated by Utility Theory

Henryk Gzyl [1,*,†] and Silvia Mayoral [2,†]

[1] Centro de Finanzas, Instituto de Estudios Superiores de Administración (IESA), Caracas 1010, Venezuela
[2] Depto de Economía de la Empresa, Charles III University of Madrid (UC3M), Getafe 28903, Spain; smayoral@emp.uc3m.es
* Correspondence: henryk.gzyl@iesa.edu.ve; Tel.: +58-212-555-4385
† These authors contributed equally to this work.

Abstract: Here, we consider the following inverse problem: Determination of an increasing continuous function $U(x)$ on an interval $[a, b]$ from the knowledge of the integrals $\int U(x)dF_{X_i}(x) = \pi_i$ where the X_i are random variables taking values on $[a, b]$ and π_i are given numbers. This is a linear integral equation with discrete data, which can be transformed into a generalized moment problem when $U(x)$ is supposed to have a positive derivative, and it becomes a classical interpolation problem if the X_i are deterministic. In some cases, e.g., in utility theory in economics, natural growth and convexity constraints are required on the function, which makes the inverse problem more interesting. Not only that, the data may be provided in intervals and/or measured up to an additive error. It is the purpose of this work to show how the standard method of maximum entropy, as well as the method of maximum entropy in the mean, provides an efficient method to deal with these problems.

Keywords: inverse problems; interpolation problems; uncertain data; maximum entropy; utility function

1. Introduction

As stated above, in this work we are interested in applying maxentropic techniques to solve integral equations in which growth and convexity constraints may be imposed on the solution. In addition, we can suppose that the data may be given up to an interval and/or be contaminated with additive errors.

Since utility theory motivates and provides a natural setup for many of the questions that we are interested in, we shall recall some very basic definitions, and now and then use the name of utility for the function that we are determining by solving the system of integral equations.

Utility theory forms the basis of both the theoretical and applied analysis to understand the process of decision making under uncertainty (see Eeckhoudt et al. [1] and Luenberger [2] for an introduction to the subject, and also the paper by Herstein and Milnor [3], in which they show how to relate preferences to utility functions).

A utility function is an increasing, continuous function $U : [a, b] \to \mathbb{R}$, which is used by a decision maker to rank random payoffs modeled by random variable taking values in $[a, b]$. A decision maker who with utility function U, is confronted with the need to choose among a variety of risky ventures or bets with random payoff X, will prefer the bet that has the largest expected utility $E[U(X)]$. To model risk aversion or risk proneness, convexity constraints are imposed upon U.

An important concept which relates the determination of a utility function to an interpolation problem is that of the *certainty equivalent*. In its simplest form, it is the quantity ξ which an investor is willing to accept instead of the outcome of a random bet, that is, it is the wealth that has the same

utility as the expected utility of a random bet X. In other words, for a given X, it is the number ξ such that:

$$U(\xi) = E[U(X)].$$

In order to display the relationship of the different problems among themselves, we organize this section as follows. First, we shall explain how the integral equation can be regarded either as generalized moment problem or an interpolation problem, then explain how the problem is to be stated when the data is provided up to an interval or when there are measurement errors. After that we shall briefly review some of the methods that have been used to solve some of those problems, and close the section describing the remainder of the paper.

1.1. Problem Statements: The Certain Data Case

Generic problem statement

Determination of an increasing and continuous function $U : [a, b] \to \mathbb{R}$, having some specified convexity properties, such that the left hand side of:

$$\pi_i = E[U(X_i)] = \int_a^b U(x) dF_i(x), \quad \text{for } i = 1, ..., M \tag{1}$$

is known for a finite collection of random variables $X_1, ..., X_M$ taking values in $[a, b]$, F_i denote their cumulative distribution functions, and π_i stands for the expected utility of each random bet X_i.

Problem 1 (The generalized moment problem). *An interesting twist on (1) is obtained if we suppose that $U(x)$ is differentiable, except perhaps at a countable set of points and suppose that $U(a)$ and $U(b)$ are known. Then we can scale U in such a way that (using U again to denote the scaled function) $U(a) = 0$ and $U(b) = 1$. Then, a simple integration by parts shows that:*

$$\pi_i = E[U(X_i)] = \int_a^b (1 - F_i(x)) u(x) dx = \int_a^b h_i(x) u(x) dx, \quad \text{for } i = 1, ..., M. \tag{2}$$

where $u(x) = dU(x)/dx$ is positive due to the fact that we are considering U to be increasing. Also $h_i(x) = (1 - F_i(x))$. Rewritten that way, (1) becomes a generalized moment problem consisting of determining the density $u(x)$ from the knowledge of the generalized moments π_i.

When the certainty equivalents ξ_i of the X_i are known, we obtain a relationship between the generalized moment problem and an interpolation problem by setting:

$$\pi_i = U(\xi_i) = \int_a^b u(x) h_i(x) dx \quad i = 1, ..., M \tag{3}$$

where now $h_i(x) = I_{[a, \xi_i]}(x)$ is the indicator function of the interval $[a, \xi_i]$, and $\{\xi_i; i = 1, ..., M\}$ is a given collection of points. Therefore, for an appropriately scaled increasing and differentiable function, solving the integral equation system is related to a generalized moment problem and to an interpolation problem. Notice as well that if the X_i in Equation (1) were certain, then (2) becomes the interpolation Problem (3).

Problem 2 (Interpolation problem). *A direct relationship between (1) and an interpolation problem comes up if we suppose that the certainty equivalent ξ_i of the X_i as well as their expected utilities π_i are known. In this case (and at this stage we do not need differentiability assumptions on the function $U(x)$) Problem (1) becomes: Determination of a continuous increasing function $U[a, b] \to \mathbb{R}$ such that:*

$$U(\xi_i) = \pi_i, \quad \text{for } i = 1, ..., M. \tag{4}$$

where the π_i are supposed to be the results of the computation of the expected value in (1). So, for this choice of random bets, the interpolation Problems (3) and the moment Problem (1) are related, except that in the former case we search for U directly, whereas in the latter we search for $u = U'$. We mention that the existence of the certainty equivalents is clear from the fact that U is continuous and increasing and that $U(a) \leq E[U(X_i)] \leq U(b)$. When considering the application to utility functions, the ξ_i are supposed to be elicited from the investor or decision maker.

1.2. Problem Statements: The Uncertain Data Case

A first extension of (1) corresponds to the case in which the data is not a single number for each X_i but an interval which contains the true number. The uncertainty may come from an uncertainty about the probability distribution of the X_i or from the data collection process. To motivate from the point of view of utility theory, we may think that all that the decision maker is able to provide is a range for the utility of a given wealth.

Problem 3 (Expected values are uncertain)**.** *Determination of an increasing and continuous function $U : [a, b] \to \mathbb{R}$, having some not yet specified smoothness or convexity properties, such that:*

$$E[U(X_i)] = \int_a^b U(x)dF_i(x) \in I_i, \quad for \ \ i = 1, ..., M. \tag{5}$$

for a finite collection of random variables $X_1, ..., X_M$, where the $I_i = [c_i, d_i], i = 1, ..., M$ are given intervals.

Problem 4 (Interpolation with uncertain data)**.** *When the X_i are deterministic and concentrated at ξ_i, or when the certainty equivalents of the X_i are known, we obtain an interpolation problem with uncertain data. The extended version of Problem 2 becomes: determination of a continuous, increasing function $U : [a, b] \to \mathbb{R}$, perhaps having some specified smoothness and convexity properties, such that:*

$$U(\xi_i) \in I_i \ \ i = 1, ..., M, \tag{6}$$

the notation being as above. Again, observe that when the values $U(a)$ and $U(b)$ are known (or set equal to 0 and 1 by appropriate rescaling), and the function U is supposed to be differentiable, by transforming Problem (5) as in (2) it becomes a generalized moment problem with uncertain data.

1.3. Problem Statements: The Case of Errors in the Data

The interpretation of this setup is a bit different. In this case we suppose that we are given values for either the utility of some possible wealths or the expected utility of some random payoffs. However, it is known that the computation of the reported number involves approximations or round off errors. The statement of the problem to solve in this case, is similar to, yet different from the two cases considered above. Consider:

Problem 5 (Expected utility computed with errors)**.** *Determination of an increasing and continuous function $U : [a, b] \to \mathbb{R}$, perhaps having some specified convexity properties, such that:*

$$E[U(X_i)] + \epsilon_i = \int_a^b U(x)dF_i(x) + \epsilon_i = \pi_i \ \ for \ i = 1, ..., M. \tag{7}$$

for a finite collection of random variables $X_1, ..., X_M$. The numbers $\pi_1, ..., \pi_M$, are known up to an additive error $\epsilon_i, i = 1, ..., M$ that has to be estimated as well. A model for the "measurement" error is available to the modeler.

Problem 6 (Utility function computed with errors). *Determination of a continuous and increasing function* $U : [a, b] \rightarrow \mathbb{R}$, *perhaps having some specified convexity properties, such that:*

$$U(\xi_i) + \epsilon_i = \pi_i \quad for \ \ i = 1, ..., M. \tag{8}$$

for a set of given numbers $\pi_1, ..., \pi_M$, *where, again, the* $\epsilon_i, i = 1, ..., M$ *are to be estimated as well. Also, a model for the "measurement" error is available to the modeler.*

1.4. Comments about the Solutions to the Generalized Moment and Interpolation Problems

We begin by saying that solving Problem (3) by means of maxentropic methods is not new. The papers by Abbas [4–6] where it seems to have been applied for the first time, and later on Abbas (2006), Darooneh [7], Pires et al. [8] and Pires et al. [9], describe different maxentropic approaches to the problems and contain references to previous work on the specific applications to utility theory. Actually, the transformation of problems (1) into (2) was considered in Abbas' papers. But none of these papers considers the case of data up to an interval or the case of errors in the data. In this regard we mention the work by Abbas [10]. Referring to the finite dimensional version of (1), he considers the possibility of prescribing ranges for the utility function or for the probability distributions to be determined.

There are many ways of solving interpolation problems like (4), from linear spline interpolation to polynomial interpolation. For interpolation methods in which convexity is a constraint, consider Mettke [11], Lai [12] and Visnawathan et al. [13]. Besides that, there is the parametric approach, which assumes that U has a known form, with shape and scale parameters that are to be fitted by lest squares. See Luenberger [2] for this approach.

We should mention that the direct maxentropic solution to Problem (3) leads to a piecewise constant $u(x)$ which in turn leads to a piecewise linear reconstruction of U.

Let us mention that there is another interesting type of problem, which in some sense is a dynamic version of the problem that we consider here. Instead of prescribing the expected values of a few random variables, one may prescribe a curve solving some optimality criterion, and then obtain a differential equation for U. In utility theory, this corresponds to determining a utility function given an optimal investment and consumption policy of an agent during a certain time period. Consider for example, Dybvig and Rogers [14] or Cox, Hobson and Obłój [15] and the references cited therein. The application of the standard maxentropic methods to economics and finance is anything but new. Consider the papers by Buchen and Kelly [16], Stutzer [17], Gulko [18] and see the review by Zhong et al. [19] for much more.

Originally, the standard method of maximum entropy was proposed in Jaynes [20] to solve problems such as (3). This method admits extensions to deal with the cases of data in intervals or errors in the data (see for example Gomes-Gonçalves et al. [21]). However, the method also admits another extension to deal with problems like (1) in which the unknown is not a density, but a function which may have any sign and not necessarily integrate to 1. Such extension seems to have been originally proposed in Navaza [22], the mathematics of which were developed by Dacunha-Castelle and Gamboa [23]. The method of maximum entropy in the mean has made its way into econometrics. See Golan et al. [24] for a review cf previous work and applications, Golan and Gzyl [25] for a unification of notation, and Gzyl and Velásquez [26] for detailed development. For applications to problems such as (1), (5), (7), see Gzyl and Mayoral [27,28]. To conclude the short digression about maximum entropy on the mean, we mention here that it is much more general and flexible than the standard method of maximum entropy, and contains a particular case (see Gzyl and ter Horst [29]). Such extension is a good tool for solving problems such as (3), (6), (8).

1.5. The Organization of the Paper

　　Having described the various versions of an integral equations problem to determine a function *U* that has to satisfy some collateral conditions, as well as two possible extensions of the problem, we mention here that the aim of this work is to present a method that solves all of these problems as variations on the theme of a single technique based on the method of maximum entropy.

　　The remainder of the paper is organized as follows: In Section 2 we shall see how to transform the integral equation problem for $U(x)$ into an integral equation for $U''(x)$. This covers the case in which the original function is required to be convex or concave. This requirement determines a signature constraint imposed on $U''(x)$. As the integral equation (for $U''(X)$) has to be solved numerically, in a short Section 3 we show how to discretize the integral equation problem in order to apply the method of maximum entropy in the mean to solve this case.

　　In Section 4 we collect the basics results on the different variations and extensions of the maximum entropy method. We begin with the standard method of maximum entropy which is a stepping stone for the method of maximum entropy in the mean, end then explain this last method.

　　Section 5 is devoted to numerical examples built upon the examples considered by Abbas (2006) and Luenberger (1998). These were chosen as representative of the two different approaches (that is, regarding the interpolation problem as a generalized moment problem or an interpolation problem). To consider extensions, in each case we either add noise to the data or consider that the data is given up to some small interval about the original data. We modify the data that they provided to examine the different possible variations on the theme. We close with a few final comments.

2. Incorporation of Convexity Constraints

　　Suppose that $U(x) : [a, b] \rightarrow \mathbb{R}$ is differentiable twice, with a bounded, measurable derivative. A simple integration by parts yields the identity:

$$U(x) = U(a) - aU'(a) + xU'(a) + \int_a^x (x - t) U''(t)dt =$$
$$U(a) - aU'(a) + xU'(a) + \int_a^b (x - t)^+ U''(t)dt.$$

　　This transforms the interpolation problem into an integral equation. At this point, we mention that even though the knowledge of the values $U(a)$ and $U(b)$ may be assumed in applications, the value of $U'(a)$ has to be estimated.

　　If the function *U* is required to be concave, a lower bound is easy to obtain. Since for $x \in [a, b]$, $U(x)$ is above the line joining $(a, U(a))$ to $(b, U(b))$, clearly $U'(a) \geq max\{U(x_i) - U(a)/(x_i - a)\}$. However, at the end, either its value or a range for it, has to be supplied by the modeler. We mention as well that the uncertainty in the specification of the derivative $U'(a)$ introduces an uncertainty in the specification of the datum that we call *m* a few lines below, which leads us to a problem with data in intervals or a problem with errors in the data.

　　Observe as well that, if $U(x)$ is required to be convex, then $U''(x)$ has to be positive, whereas if $U(x)$ is to be concave, then U'' must be negative. To be specific, with applications to utility theory in mind, we shall consider the risk-averse decision maker and require $U''(x)$ to be positive. To simplify the notation, shall write $v(x) = -U''(x)$ and recast Problem (3) as: finding a positive function $v(x)$ such that:

$$\int_a^b k(x_i, t)v(t)dt = m_i \ \ \text{where} \ \ k(x, t) = (x - t)^+, \ \ i = 1, ..., M. \tag{9}$$

where $m_i = x_i U'(a) - U(x_i) - aU'(a) + U(1)$. Note that the right hand side of this identity is positive. Coming back to the statement of Problem 1, note that under the same conditions on *U* mentioned above, substituting $U'(x) = U'(a) + \int_a^t U''(y)dy$ in (2) and using de fact that $\int_a^b (1 - F(x))dx = E[X_i] - a$, we obtain after exchanging integrals:

$$E[U(X_i)] = U(a) - aU'(a) + U'(a)E[X_i] + \int_a^b U''(t) \left(\int_t^b (1 - F_i(s)) ds \right) dt. \tag{10}$$

Again, note that when X_i is certain, say $P(X_i = \xi_i) = 1$ for some $\xi_i \in [a,b]$, then $E[X_i] = \xi_i$, $1 - F(t) = I_{[0,\xi_i]}(t)$ and $\int_t^b (1 - F(s)) ds = (\xi_i - t)^+$. Thus the interpolation problem with convexity constraints becomes an integral equation problem like:

$$U(\xi_i) = m_i + \int_a^b v(t)(\xi_i - t)^+ dt \tag{11}$$

in which the specification of the data depends not only on $U(\xi_i)$ but on m_i as well.

To make all these problems look formally alike, write $k_i(t) = k(x_i, t)$ and rewrite (9) as:

$$\int_a^b k_i(t)v(t)dt = m_i, \quad i = 1, ..., M. \tag{12}$$

and when the data is uncertain, we would need to solve:

$$\int_a^b k_i(t)v(t)dt \in [c_i, d_i], \quad i = 1, ..., M. \tag{13}$$

Clearly, all the problems described in Section 1 lead to our wanting to solve the same generalized moment problems like (12) and (13), with boundedness constraints imposed on the solution.

3. Discretization of the Problems

Since we want to solve the problems numerically, we must discretize the integral equations introduced in the previous sections. Let us describe the generic problem: Determine a continuous function $v : [a,b] \to [0,B]$ (or $v : [a,b] \to [0,\infty)$) that satisfies:

$$\int_a^b k_i(t)v(t)dt \in [c_i, d_i], \quad i = 1, ..., M. \tag{14}$$

Notice that when $c_i = d_i$ for $i = 1, ..., M$, we are reduced to case of solving a standard generalized moment problem. Denote by $t_j = a + (j-1)/N$ with $j = 1, ..., N$ the left end points of a partition of $[a,b]$ into N sub-intervals. Define the vector $v \in \mathbb{R}^N$ to have components $v_j = v(t_j)$. Define the matrix A by its components $A_{i,j}$ constructed from the datum $k_i(t)$ as:

$$A_{i,j} = \frac{1}{N}k_i(t_j).$$

Define $\mathcal{D} = \prod_{i=1}^M [c_i, d_i]$ to be the data region and $\mathcal{K} = [0,B]^N$ (or $\mathcal{K} = [0,\infty)^N$) to be the constraint set, that is the set in which the solution to the discretized problem has to lie. With this notation, the discrete version of Problem (14) becomes:

$$\text{Find } v \in \mathcal{K} \subset \mathbb{R}^N \text{ such that } Av \in \mathcal{D} \subset \mathbb{R}^M. \tag{15}$$

which is a standard linear inverse problem with (non-linear) convex constraints. As said, the method of maximum entropy in the mean handles such problems in a systematic fashion.

4. The Maximum Entropy Methods

The results presented below are not new, and are included for completeness. It may be skipped by readers familiar with the methodology. The complete proofs of the statements about convex duality are provided in Borwein and Lewis [30] and in Gzyl and Velásquez [26].

This section is divided into three subsections. In the first one we review the standard method of maximum entropy (SME), then the second its extension (SMEE) to incorporate uncertain data. As we explained, these will be applied to the first two problems and to their extension when the data is uncertain or has errors. Then we review the method of maximum entropy in the mean (MEM) and its extension to the case of uncertain data. We shall apply this technique to the case in which one wants to determine the second derivative of U from the data about preferences.

4.1. The Standard Maximum Entropy Method (SME)

The standard maximum entropy method (SME) is the name of a variational procedure to solve the (inverse) problem consisting of finding a probability density $u(t)$ (on $[0,1]$) in our case, satisfying the following integral constraints:

$$\int_0^1 k_i(t)u(t)dt = m_i \qquad \text{for} \qquad i = 0, 1, ..., M. \tag{16}$$

The notation follows that introduced during the generic statement of Problems 1 to 6. The material presented below is standard, and here we recall the results that we use for the numerical computations. We set $k_0 \equiv 1$ and $\mu_0 = 1$ since we require $u(t)$ to be a density function. It actually takes a standard computation to see that when the problem has a solution it is of the type:

$$u^*(t) = \exp\left(-\sum_{i=0}^M \lambda_i^* k_i(t)\right) \tag{17}$$

in which the number of moments M appears explicitly. It is usually customary to write $e^{-\lambda_0^*} = Z(\boldsymbol{\lambda}^*)^{-1}$, where $\boldsymbol{\lambda}^* = (\lambda_1^*, ..., \lambda_M^*)$ is an $M-$dimensional vector. Clearly, the generic form of the normalization factor is given by:

$$Z(\boldsymbol{\lambda}) = \int_0^1 e^{-\sum_{i=1}^M \lambda_i k_i(t)}dt. \tag{18}$$

With this notation the generic form of the solution can be rewritten as:

$$u^*(t) = \frac{1}{Z(\boldsymbol{\lambda}^*)}e^{-\sum_{i=1}^M \lambda_i^* k_i(t)} = e^{-\sum_{i=0}^M \lambda_i^* k_i(t)}. \tag{19}$$

The vector $\boldsymbol{\lambda}^*$ can is found minimizing the dual entropy:

$$\Sigma(\boldsymbol{\lambda}, \boldsymbol{m}) = \ln Z(\boldsymbol{\lambda}) + <\boldsymbol{\lambda}, \boldsymbol{m}> \tag{20}$$

where $< \mathbf{a}, \mathbf{b} >$ denotes the standard Euclidean scalar product and m is the $M-$vector with components m_i, and obviously, the dependence of $u^*(t)$ on m is through $\boldsymbol{\lambda}^*$. A simple calculation shows the first order condition for $\boldsymbol{\lambda}^*$ to be a minimizer of $\Sigma(\boldsymbol{\lambda}, \boldsymbol{m})$, which is equivalent to saying that (19) satisfies (16).

4.2. The Standard Maximum Entropy for Data in Ranges (SMEE)

Here we present an extension of the method described above to solve the (inverse) problem consisting of finding a probability density $u(t)$ (on $[0,1]$ in this case), satisfying the following integral constraints:

$$\int_0^1 k_i(t)u(t)dt \in I_i \qquad \text{for} \qquad i = 0, ..., M. \tag{21}$$

where the interval $I_i = [c_i, d_i]$, around the true but unknown m_i depends on the undecided decision maker. For $i = 0$ only we set $I_0 = \{1\}$ since for $k_0 \equiv 1$ we have $m_0 = 1$ to take care of the natural normalization requirement on $u(t)$.

Let us put $\mathcal{D} = I_1 \times ... \times I_M$. The idea behind the SMEE is to solve a SME problem for each $\mathbf{m} \in \mathcal{D}$, and then maximize over the entropy of these solutions. That eventually leads to the need of minimization of the following dual entropy function:

$$\Sigma(\boldsymbol{\lambda}) = \ln(Z(\boldsymbol{\lambda})) + \delta_{\mathbf{C}}^*(\boldsymbol{\lambda})$$

where:

$$\delta_{\mathcal{D}}^*(\boldsymbol{\lambda}) \equiv \sup\{< \boldsymbol{\lambda}, \mathbf{c} > |\mathbf{m} \in \mathcal{D}\} = \sum_{i=1}^M \frac{(d_i - c_i)|\lambda_i|}{2} + \sum_{i=1}^M \frac{(d_i + c_i)\lambda_i}{2}$$

is the Fenchel–Legendre dual of the indicator $\delta_{\mathcal{D}}$ of the convex set $\mathcal{D}$, (see Borwein and Lewis (2000) for these matters). To obtain this result, just notice that $[-1, 1]^M$ can be mapped onto $\mathcal{D}$ by means of affine mapping, and:

$$\sup\{< \boldsymbol{\lambda}, \mathbf{x} > |\mathbf{x} \in [-1, 1]^M\} = \|\boldsymbol{\lambda}\|.$$

As a first step towards describing the solution, we state the following result, the proof of which can be seen in Borwein and Lewis [30].

Lemma 1. *With the notations introduced above, set*

$$\Sigma(\boldsymbol{\lambda}) = \ln(Z(\boldsymbol{\lambda})) + \delta_{\mathcal{D}}^*(\boldsymbol{\lambda})$$

Then $\Sigma(\boldsymbol{\lambda})$ is strictly convex in $\boldsymbol{\lambda}$.

If the minimum of $\Sigma(\boldsymbol{\lambda})$ is reached at a point $\boldsymbol{\lambda}^*$ in the interior of the set $\{\boldsymbol{\lambda} \in \mathbb{R}^M | Z(\boldsymbol{\lambda}) < \infty\}$, then maxentropic solution to the inverse problem (21) is given by:

$$u^*(y) = \frac{1}{Z(\boldsymbol{\lambda}^*)} e^{-\sum_{k=1}^M \lambda_i^* k_i(t)} = e^{-\sum_{i=0}^M \lambda_i^* k_i(t)}. \tag{22}$$

Since the numerical minimization of $\Sigma(\boldsymbol{\lambda})$ is necessary for the determination of the optimal $\boldsymbol{\lambda}$, the following computation is worth keeping in mind. Observe that $\partial\delta_{\mathcal{D}}^*(\boldsymbol{\lambda})/\partial\lambda_i$ is defined except at $\lambda_i = 0$, where it is sub-differentiable (see Borwein and Lewis (2000)). Actually,

$$\partial\delta_{\mathcal{D}}^*(\boldsymbol{\lambda})/\partial\lambda_i = d_i \text{ when } \lambda_i > 0,$$
$$\partial\delta_{\mathcal{D}}^*(\boldsymbol{\lambda})/\partial\lambda_i = c_i \text{ when } \lambda_i < 0,$$
$$\partial\delta_{\mathcal{D}}^*(\boldsymbol{\lambda})/\partial\lambda_i \in (c_i, d_i) \text{ when } \lambda_i = 0 \tag{23}$$

Keeping the above computations in mind, it is easy to verify that the first order condition for $\boldsymbol{\lambda}^*$ to be a minimizer yields.

$$\int_0^1 k_i(t) u^*(t) dt \in [c_i, d_i], \text{ for } i = 1, ..., M.$$

4.3. The Method of Maximum Entropy in the Mean

In this section we come back to the problem of finding an increasing, continuous function $U : [a, b] \to \mathbb{R}$, with given boundary behavior and specified convexity constraints. Even though the method that we shall implement can take care of constraints like $-B_1(t) \le U''(t) \le B_2(t)$, for given functions B_1 and B_2, we shall content ourselves with one of two possible cases: $0 \le U''(t) \le B$ for some finite B or the case $0 \le U''(t) < \infty$, corresponding to the cases in which the second derivative of U is known to be bounded or not.

Recall that to solve the discrete version of Problem (15) we need to solve the linear system:

$$Av \in \mathcal{D} \quad \text{where} \quad v \in \mathcal{K}$$

where $\mathcal{K}$ will be either $[0, B]^N$ or $[0, \infty)^N$, and $\mathcal{D} = \prod_{i=1}^{M}[c_i, d_i]$. Let us explain how to solve it using MEM.

4.3.1. The Intuition Behind MEM

Maximum entropy in the mean is the name of a technique which allows us to transform an algebraic problem into a non-linear optimization problem of smaller dimension. The standard presentation of the method is explained in Gzyl and Velásquez [26]. Let us devote a few lines to explain the intuition behind it.

Instead of solving (15) we search for a probability P on $(\mathcal{K}, \mathcal{F})$, where $\mathcal{F}$ denotes the Borel $\sigma-$algebra of subsets of $\mathcal{K}$, such that $E_P[V] = v^*$ solves (15), that is $Av^* \in \mathcal{D}$. Here $V : \mathcal{K} \to \mathcal{K}$ is the identity (or coordinate) map on $\mathcal{K}$. To find P one proposes any interesting reference $\sigma-$finite measure or probability Q, such that the closure of the convex hull of its support is $\mathcal{K}$, and then one uses a maxentropic procedure to determine a positive density $\rho(v)$ of P with respect to Q, such that:

$$AE_P[V] = A \int_{\mathcal{K}} v\rho(v)dQ(v) \in \mathcal{D}. \tag{24}$$

It only remains to be added that it is by finding $\rho(v)$ satisfying (24) that one solves a maximum entropy problem for each $y \in \mathcal{D}$, and then "best" density is determined as the one that maximizes the corresponding entropies.

4.3.2. MEM for the Problem with Preassigned Bounds

This part covers both the initial step for MEM with certain data as well as MEM for the problem with uncertain data. There are two cases to consider, depending on whether we want the unknown function to have preassigned bounds or not. The procedure is the same, but the constraint space is not.

Case 1: $0 \leq v(t) \leq B$

This is certainly the simple possible case, but it contains all the ingredients for the most general case. As any point in $[0, B]$ can be written as a convex combination of 0 and B, we choose as reference measure on $\mathcal{K} = [0, B]^N$ the measure:

$$dQ(v) = \prod_{j=1}^{N} \left(\epsilon_0(dv_j) + \epsilon_B(dv_j) \right)$$

where $\epsilon_B(dv_j)$ denotes the unit point mass (Dirac's measure) at B in the $-$th direction. Since any $P << Q$ may be written as:

$$dP(v) = \prod_{j=1}^{N} \left((1 - p_j)\epsilon_0(dv_j) + p_j\epsilon_B(dv_j) \right)$$

the constraint Equation (24) become:

$$AE_P[V] = B \sum_{j=1}^{N} A_{i,j}p_j = y_i, \quad i = 1, ..., M, \tag{25}$$

where $y_i \in [c_i, d_i]$ is any point. To find the p_j's we introduce the usual entropy:

$$S(P) = -\int_{\mathcal{K}} \frac{dP}{dQ}(v) \ln\left(\frac{dP}{dQ}(v)\right) dQ(v) = -\sum_{j=1}^{N} (1 - p_j) \ln(1 - p_j) + p_j \ln p_j.$$

The standard argument leads to the representation:

$$p_j^* = \frac{e^{-B \sum_{i=1}^{M} A_{i,j} \lambda_i^*}}{Z(\lambda)} \quad \Rightarrow v_j^* = B p_j, \tag{26}$$

where the identification in the last step can be read from (25). The normalization factor $Z(\lambda)$ is given by:

$$Z(\lambda) = \sum_{j=1}^{N} e^{-B \sum_{i=1}^{M} A_{i,j} \lambda_i} \quad \lambda \in \mathbb{R}^M,$$

and finally, the λ_i^* are to be found minimizing the dual entropy function:

$$\Sigma(\lambda, \mathbf{y}) = \ln(Z(\lambda)) + \langle \lambda, \mathbf{y} \rangle.$$

Case 2: No bounds preassigned

This time $\mathcal{K} = [0, \infty)$, and the rest of the notation has the same meaning as in the first case. A possible candidate for reference measure is a product of Poisson measures (of parameter $= 1$). This is so because they are supported by the non-negative integers and any point in the line can be obtained by a convex combination of integers. So, we consider:

$$dQ(v) = \prod_{j=1}^{N} \left(\sum_{n=0}^{\infty} \frac{1}{n!} e^{-1} \epsilon_n(dv_j) \right).$$

If we denote by $\rho(v)$ the density of the unknown measure P with respect to Q, the constraint equations look like:

$$\sum_{i,j} A E_P[V_j] = \sum_{j=1}^{N} A_{i,j} \int_{\mathcal{F}} \rho(v) v_j dQ(v) = y_i, \quad i = 1, ..., M. \tag{27}$$

Now we have to maximize:

$$S(P) = -\int_{\mathcal{K}} \rho(v) \ln \rho(v) dQ(v)$$

subject to (27), plus the normalization condition. The standard argument leads to the representation:

$$\rho^*(v) = \frac{1}{Z(\lambda)} \prod_{j=1}^{N} e^{-\sum_{i=1}^{M} \lambda_i^* A_{i,j} v_j} \Rightarrow v_j^* = \int_{\mathcal{K}} v_j dP(v) = e^{-\sum_{i=1}^{M} \lambda_i^* A_{i,j}}. \tag{28}$$

This time the normalization factor is easily computed to be:

$$Z(\lambda) = \int_{\mathcal{K}} \prod_{j=1}^{N} e^{-\sum_{i=1}^{M} \lambda_i^* A_{i,j} v_j} dQ(v) = e^{-\sum_{j=1}^{N} \left(1 - \exp\left(-\sum_{i=1}^{M} \lambda_i^* A_{i,j}\right)\right)}.$$

The vector $\boldsymbol{\lambda}^*$ is found, minimizing:

$$\Sigma(\boldsymbol{\lambda}, \mathbf{y}) = \ln\left(Z(\boldsymbol{\lambda})\right) + \langle \boldsymbol{\lambda}, \mathbf{y} \rangle = -\sum_{j=1}^{N}\left(e^{-\sum_{i=1}^{M}\lambda_i A_{i,j}} - 1\right) + \langle \boldsymbol{\lambda}, \mathbf{y} \rangle.$$

The representation mentioned in the right hand side of (28) appears when we apply the first order condition for $\boldsymbol{\lambda}^*$ to be a minimizer of $\Sigma(\boldsymbol{\lambda}, \mathbf{y})$.

4.4. Maximum Entropy in the Mean for the Case with Uncertain Data

Let us now address the problem of determining the utility function with prescribed convexity properties, when the utility function of the decision maker is known up to a range of values. We know that this problem reduces to the problem of determining $v \in \mathbb{R}^N$ such that:

$$\mathbf{Av} \in \mathcal{D} \quad \mathbf{v} \in \mathcal{K}.$$

The rationale to solve this problem is similar to that used in the SME with uncertain data: Solve a MEM problem for each $\mathbf{y} \in \mathcal{D}$, and then maximize over the entropies of the resulting maxentropic densities.

That leads to the need to determine the minimizer of the dual entropy:

$$\Sigma(\boldsymbol{\lambda}) = \ln Z(\boldsymbol{\lambda}) + \delta_{\mathcal{D}}^*(\boldsymbol{\lambda})$$

over $\mathbb{R}^M$. We thus obtain a maxentropic density $\rho^*(\mathbf{v})$ of over $\mathcal{K}$ such that the mean:

$$\mathbf{v}^* = \int_{\mathcal{K}} \mathbf{v}\rho^*(\mathbf{v})dQ(\mathbf{v})$$

is the solution to $\mathbf{Av} \in \mathcal{D} \quad \mathbf{v} \in \mathcal{K}$. To go beyond these generalities, we must specify $\mathcal{K}$ and Q.

4.4.1. MEM for the Problem with Uncertain Data and Preassigned Bound on the Second Derivative

We consider the same two cases as before, according to whether the second derivative is bounded or not, but in both cases it is required to be positive. In both cases the set up and the arguments are similar to those used in Section 4.2 for the SME method when the data is uncertain. The representation of the density and the normalization factor are the same. All that changes is the dual function to be minimized to determine the optimal $\boldsymbol{\lambda}^*$, which this time is:

$$\Sigma(\boldsymbol{\lambda}) = \ln Z(\boldsymbol{\lambda}) + \delta_{\mathcal{D}}^*(\boldsymbol{\lambda})$$

where the specification of the domain $\mathcal{D}$ is chosen according the un-decidedness of the decision maker.

Case 1: $0 \le v(t) \le B$

In this case the normalization factor found above is:

$$Z(\boldsymbol{\lambda}) = \sum_{j=1}^{N} e^{-B\sum_{i=1}^{M}A_{i,j}\lambda_i} \quad \boldsymbol{\lambda} \in \mathbb{R}^M,$$

Therefore, now we have to find the minimizer of:

$$\Sigma(\boldsymbol{\lambda}) = \ln\left(\sum_{j=1}^{N} e^{-B\sum_{i=1}^{M}A_{i,j}\lambda_i}\right) + \sum_{i=1}^{M}\frac{d_i - c_i}{2}|\lambda_i| + \sum_{i=1}^{M}\frac{d_i + c_i}{2}\lambda_i.$$

Also, once the optimal λ^* is found, the maxenntropic solution to our problem is:

$$v_j^* = B\frac{e^{-B\sum_{i=1}^{M} A_{i,j}\lambda_i^*}}{Z(\lambda^*)}. \tag{29}$$

Note that the solution looks similar that of the case in which the data is certain, except that this time the dual entropy is different. Note as well that when $c_i = d_i = y_i$, the two solutions to the two problems become identical.

Case 2: Positive $v(t)$

As said, the setup described in Section 4.2 can be transported verbatim. The normalization factor for this case is:

$$Z(\lambda) = \int_{\mathcal{K}} \prod_{j=1}^{N} e^{-\sum_{i=1}^{M}\lambda_i A_{i,j}v_j}dQ(v) = e^{-\sum_{j=1}^{N}\left(1-\exp(-\sum_{i=1}^{M}\lambda_i^* A_{i,j})\right)}.$$

and therefore, to determine the optimal λ^*, the function to be minimized in this case is:

$$\Sigma(\lambda) = -\sum_{j=1}^{N}\left(1 - \exp(-\sum_{i=1}^{M}\lambda_i A_{i,j})\right) + \sum_{i=1}^{M}\frac{d_i - c_i}{2}|\lambda_i| + \sum_{i=1}^{M}\frac{d_i + c_i}{2}\lambda_i$$

This time, once the minimizer λ^* has been found, the maxentropic solution to the problem of determining the solution to (15) is:

$$v_j^* = \exp(-\sum_{i=1}^{M}\lambda_i^* A_{i,j}) \tag{30}$$

4.5. Case with Errors in the Data and Uncertainty in the Data

This case extends Problem 6, and it was motivated by the following considerations. Notice that in the statement of (9), on one hand, the value of $U(\xi_i)$ may only be known up to an interval, and on the other, when transforming the problem into an integral equation for the second derivative of U, and error may be introduced in the estimate of $U'(a)$.

Thus we are led to extend the statement of (24) to:

$$AE_P[\mathbf{V}] + \eta = A\int_{[0,\infty)^N}\mathbf{v}\rho(\mathbf{v})dQ(\mathbf{v}) + \eta \in \mathcal{D}. \tag{31}$$

in which $\eta \in \prod_{i=1}^{M}[-\delta_k, \delta_k]$ is a vector whose components lie in intervals containing the data points $m_i = x_iU'(a) - U(x_i) - aU'(a) + U(1)$. The uncertainty in the values of the function U and the δ_k capture the uncertainty due to errors in the estimation of $U'(a)$ say. Stated this way, the problem clearly extends the statement of Problem 6, in which the decision maker may be certain about his choice, but he misestimates the m_i.

To solve this problem the intuition about MEM carries on to this case. All we have to add is that we also want to think of the η_k as expected values with respect to an probability law that must also be determined.

So as not to introduce much more symbols, denote by $A_e = [A, I]$, where I is the $M \times M$ identity matrix and $x^t = (v, y)$. To simplify our lives we introduce $\mathcal{K}_e = [0,\infty)^N \times \prod_{i=1}^{N}[-\delta,\delta]$, that is, we suppose that all errors have the same bound.

Now, (15) becomes:

$$\text{Find } x \in \mathcal{K}_e \subset \mathbb{R}^N \text{ such that } Ax \in \mathcal{D} \subset \mathbb{R}^M. \tag{32}$$

Note that when $\mathcal{D}$ reduces to a point, we are back in the case of Problem 6. Again, the first step of the routine consists of choosing a measure Q_e on $(\mathcal{K}_e, \mathcal{B}(\mathcal{K}_e))$ to play the role of reference measure. We shall chose a product measure $Q_e(dv, dy) = Q_s(dv)Q_n(dy)$ (think of the subscripts as "signal" and "noise"). The $Q_s(dv)$ is as above and for the noise we choose:

$$Q_n(dy) = \prod_{i=1}^{N} \left(\epsilon_{-\delta}(dy_i) + \epsilon_{\delta}(dy_i) \right).$$

The Laplace transform of this measure is:

$$\int_{\mathcal{C}_e} e^{-\langle \tau, x \rangle} dQ_e(dv, dy) = \prod_{j=1}^{N} e^{-(1-\exp \tau_j)} \prod_{i=1}^{M} \left(e^{-\delta \tau_{N+i}} + e^{\delta \tau_{N+i}} \right).$$

To define $Z(\lambda)$ we replace $\tau_j : 1 \leq j \leq N+M$ by $(A_e^t \lambda)_j$. The density $\rho(x)$ that maximizes the entropy is, again, given by:

$$\rho(x) = \frac{e^{-\langle \boldsymbol{\lambda}^*, A_e x \rangle}}{Z(\boldsymbol{\lambda})}$$

where λ^* is obtained maximizing the dual entropy:

$$\Sigma(\boldsymbol{\lambda}) = \ln Z(\boldsymbol{\lambda}) + \sum_{i=1}^{M} \frac{d_i - c_i}{2} |\lambda_i| + \sum_{i=1}^{M} \frac{d_i + c_i}{2} \lambda_i.$$

Bringing in the expression for $Z(\lambda)$ the last identity becomes:

$$\Sigma(\boldsymbol{\lambda}) = -\sum_{j=1}^{N} \left(1 - e^{-A^t \lambda_j} \right) + \sum_{i=1}^{M} \ln \left(e^{\delta \lambda_i} + e^{-\delta \lambda_i} \right) + \sum_{i=1}^{M} \frac{d_i - c_i}{2} |\lambda_i| + \sum_{i=1}^{M} \frac{d_i + c_i}{2} \lambda_i. \tag{33}$$

Once the $\boldsymbol{\lambda}^*$ that minimizes (33) has been determined, the representation of the (discretized) second derivative is given by an expression similar to (30), namely:

$$v_j^* = \exp(- \sum_{i=1}^{M} \lambda_i^* A_{i,j}). \tag{34}$$

Also, the estimated errors are given by:

$$\eta_i^* = \frac{-\delta e^{\delta \lambda_i^*} + \delta e^{-\delta \lambda_i^*}}{e^{\delta \lambda_i^*} + e^{-\delta \lambda_i^*}}. \tag{35}$$

5. Numerical Examples

We shall consider extended versions of the numerical problems treated by Abbas and by Luenberger. The extensions go beyond the examples in two ways: First, we include the possibility that the data is known up to a range, and second, we include convexity constraints upon the unknown.

5.1. Abbas' Example with and without Uncertainty

In Abbas (2006), the author uses the data in Table 1 and Problem (1) to determine $u(t) = U'(t)$ regarded as a probability density on $[0, 1]$. The scaling that he uses comes in handy for that interpretation. The first row in Table 1 contains a few possible wealths and in the second row we display the utility of that wealth. In the last row we list possible ranges of utility for the undecided decision maker.

In Figure 1 we display following three possible reconstructions, all using the SME method. In the continuous line we present the reconstruction corresponding to the original data. The upper curve

corresponds to the reconstruction using as inputs the upper ends points of the undecidability region, and the lower curve displays the utility corresponding to the lower end points of the undecidability region. The ordering of the utility functions is intuitively clear. This is because this approach uses the data in the third row in a trivial way, and since the utility function is increasing and interpolates a sequence of point values, the larger the values, the larger the interpolated function. Therefore, any other method using the interval data, must produce a result within those two graphs.

Table 1. Extended Abbas' data.

x	0	20	40	50	90	100
$U(x)$	0	0.32	0.57	0.67	0.95	1
Range	[0, 0.01]	[0.30, 0.34]	[0.52, 0.59]	[0.64, 0.7]	[0.92, 0.98]	[0.99, 1]
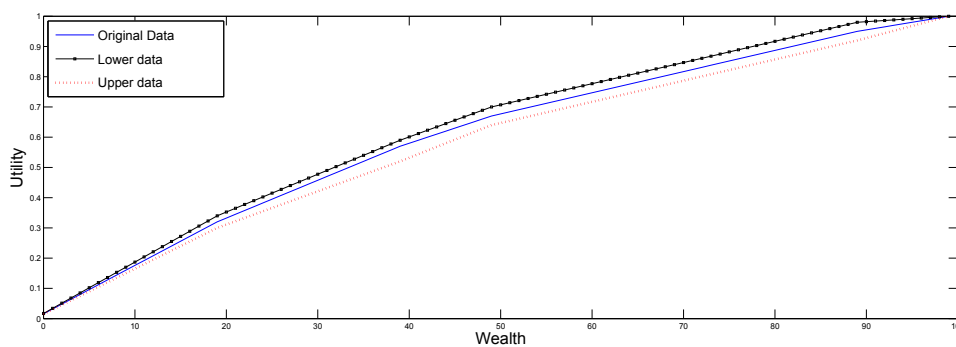
Figure 1. Utility reconstruction from data in Table 1.

To continue with the example, in Figure 2 we present in continuous line, the reconstruction according to the SME procedure for data in ranges described in Section 4.2, and we plotted the two utility functions determined by the end points of the undecidability range for comparison. All these curves are in the same plot just to visualize them together, not for comparison purposes because they solve to different, albeit related problems. Note that now, the "true" utility falls within the range determined by the utilities corresponding to the upper and lower preference values.
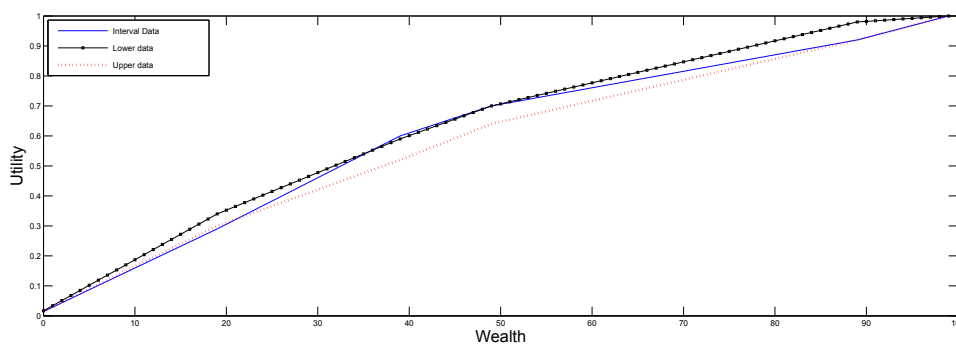
Figure 2. Utility reconstructed from data in ranges in Table 1.

5.2. Luenberger's Example with and without Uncertainty

This example is worked out in Luenberger [2] by means of a least squares minimization procedure applied to a power function that has the right convexity properties built in from the beginning. He chooses $U(x) = ax^\gamma + c$ and recovers $a = 4, c = -3$, and $\gamma = 1/2$, which renders $U(x)$ concave. The input data is presented in the original form in Table 2, and rescaled in Table 3.

Table 2. Extended Luenberger's example.

x	1	1.44	1.96	2.56	3.24	4	4.84	5.76	6.76	7.84	9
$U(x)$	1	1.8	2.6	3.24	4.2	5	5.8	6.6	7.4	8.2	9
Range	[0.09, 1.01]	[1.7, 1.9]	[2.45, 2.75]	[3.22, 3.27]	[4.15, 4.28]	[4.9, 5.1]	[5.6, 6.0]	[6.4, 6.8]	[7.2, 7.6]	[8.1, 8.35]	[8.9, 9]

**Table 3.** Standardized extended Luenberger's table.

x	1	1.44	1.96	2.56	3.24	4	4.84	5.76	6.76	7.84	9
$U(x)$	0	0.1	0.2	0.28	0.4	0.5	0.6	0.7	0.925	0.9	1
Range	[0, 0.01]	[0.08, 0.112]	[0.18, 0.22]	[0.28, 0.34]	[0.39, 0.41]	[0.49, 0.51]	[0.56, 0.6]	[0.675, 0.725]	[0.68, 0.83]	[0.89, 0.92]	[0.99, 1]

In Figure 3 we plot the original parametric curve (in dots), as well as the utility reconstructed by applying the standard maximum entropy method to the data in Table 3, the first row of which lists the certainty equivalents corresponding to the expected utilities listed in the second row. Clearly the agreement is quite good, and we did not have to guess the right parametric representation of the function.
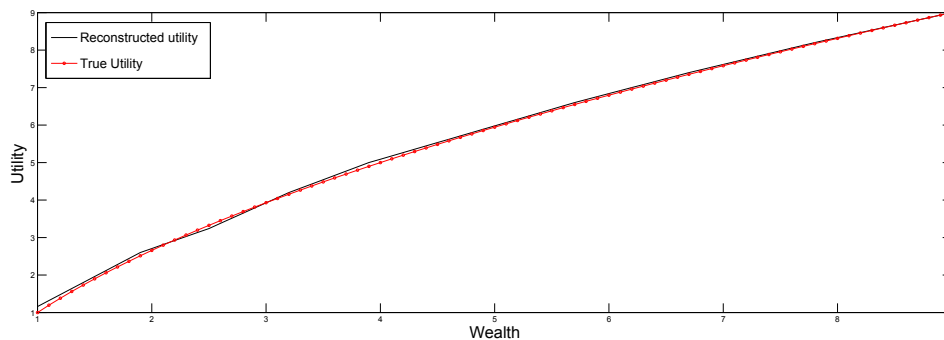
Figure 3. Reconstructed utility from point data in Table 3.

In the third row of Tables 2 and 3 we list some possible interval ranges of the utility function of an undecided decision maker (original and scaled down ranges). To show the usefulness of the maxentropic procedure when the data is given in ranges, we apply the maxentropic procedure for data in ranges explained in Section 4.2, to the data in the third row of Table 3, and display the utility function so obtained in Figure 4. Clearly, when the data is specified in ranges, it would be somewhat hard to apply a parametric fitting procedure. The maxentropic procedure produces an increasing and continuous function that stays within the specified utility range.
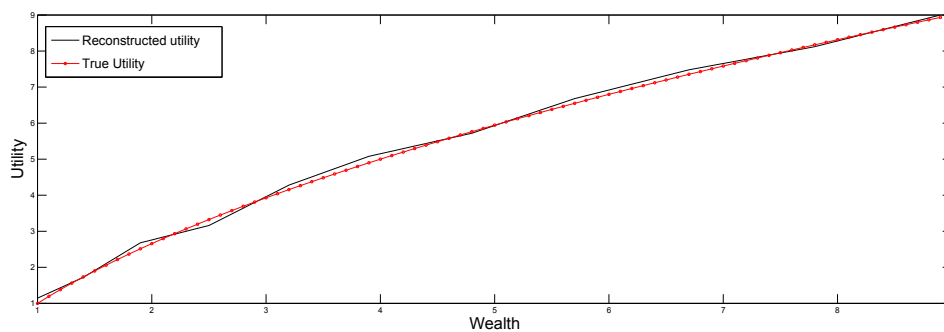
**Figure 4.** Utility reconstructed from interval data of Table 3.

Now, to consider another variation on the theme, suppose that we want to make sure that the utility function has the right concavity (or convexity) properties. Certainly, a simple plot of the vales of the utility suggests that the utility must be concave. Thus, we might decide to consider the problem analyzed in Section 2, that is to reconstruct the second derivative of the utility function by means of the method of maximum entropy in the mean described in Section 4.3. Accordingly, we have to solve (9), where the vector *m* that is specified right below that equation to be used as input for the numerical analysis is given Table 4.

Table 4. Modified Luenberger's data.

x	1	1.44	1.96	2.56	3.24	4	4.84	5.76	6.76	7.84	9
m	0	0.08	0.32	0.72	1.28	2.00	2.88	3.92	5.12	6.48	8.00

Recall that the procedure of maximum entropy in the mean is used now to determine the second derivative of the utility function, from which the function itself is determined by simple integration.

The reconstruction obtained from the data in Table 4 is displayed in Figure 5 along with the original function. Recall that the output of this procedure is the second derivative of the utility function, which is then integrated to recover the function itself. This reconstruction is very sensitive to the input data and the discretization procedure, that is why it is convenient to apply the version of MEM that allows for data in intervals as well as for errors in the data. The procedure was described in Section 4.5, and the result of applying it to the data listed in Table 5 is displayed in Figure 6 for two values of the parameter δ describing the size of the error.
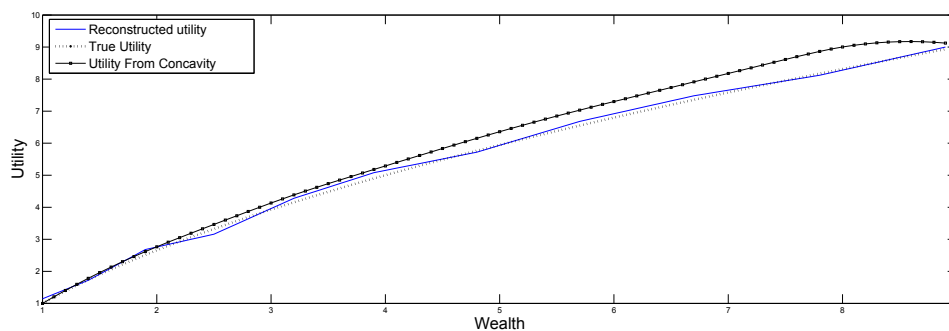
Figure 5. Utility function from reconstructed second derivatives from data in Table 4.

Table 5. Data for determining the function from reconstructed second derivative.

x	1	1.44	1.96	2.56	3.24	4	4.84	5.76	6.76	7.84	9
m	[0, 0.01]	[0.07, 0.09]	[0.31, 0.33]	[0.71, 0.73]	[1.27, 1.29]	[1.99, 2.01]	[2.87, 2.89]	[3.91, 3.93]	[5.11, 5.13]	[6.47, 6.49]	[7.99, 8.01]

The data in Table 5 corresponds to a decision maker that is slightly undecided. The two utility functions obtained from the reconstructed discretized second derivative according to the results presented in sections two and three. The displayed reconstructions correspond to $\delta = 0.01$ and $\delta = 0.09$.

In Figure 6 we present two utility functions obtained from the reconstructed discretized second derivative according to the results presented in Sections 2 and 3. The displayed reconstructions correspond to $\delta = 0.01$ and $\delta = 0.09$. We mention that for values of $\delta \geq 0.1$ the algorithm converges too slowly because the derivative of the dual entropy oscillates from one end of the reconstruction range to the other without converging to zero.
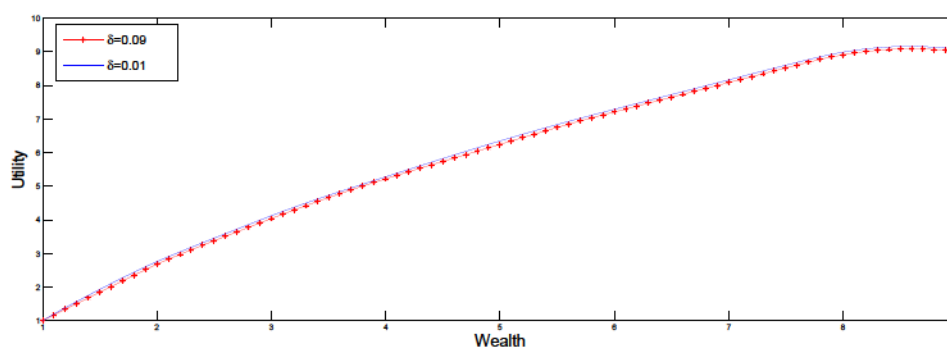
Figure 6. Utility function from reconstructed second derivatives from data in Table 5.

6. Concluding Remarks

To sum up and highlight the applications of utility theory, the problem of determining preferences from the expected value of a few random bets of from their certainty equivalents can be rephrased as an interpolation problem which can itself be transformed into a generalized moment problem. The moment problem can be extended to incorporate errors in the data and data in ranges. This would correspond to the case in which there are errors in the data of the utility function or that the utility is known up to a range. All of these problems can be solved by appropriate versions of the maximum entropy method.

Furthermore, when the risk preferences of the decision maker are known, the problem mentioned in the previous paragraph can be transformed into a problem consisting in solving a linear integral equation for a function with pre-specified signature. The method of the maximum entropy in the mean can be applied to solve that problem for us.

Author Contributions: Both authors contributed equally to the paper. Both the authors have read and approved the final manuscript.

Conflicts of Interest: The authors declare no conflict of interest.

Abbreviations

The following abbreviations are used in this manuscript:

SME   standard maximum entropy
MEM   maximum entropy in the mean
MDPI   Multidisciplinary Digital Publishing Institute
DOAJ   Directory of open access journals
TLA   Three letter acronym
LD   linear dichroism

References

1. Eeckhoudt, L.; Gollier, C.; Schlessinger, H. *Economic and Financial Decisions under Risk*; Princeton University Press: Princeton, NJ, USA, 2005.
2. Luenberger, D. *Investment Science*; Oxford University Press: Oxford, UK, 1998.
3. Herstein, I.; Milnor, J. An axiomatic approach to measurable utility. *Econometrica* **1953**, *21*, 291–297.
4. Abbas, A.E. An entropy approach for utility assignment in decision analysis. In *Bayesian Inference and Maximum Entropy Methods in Science and Engineering*; Williams, C.J., Ed.; Kluwer Academic: Dordrecht, The Netherlands, 2003; pp. 328–338.

5. Abbas, A.E. Entropy methods for univariate distributions in decision analysis. In *Bayesian Inference and Maximum Entropy Methods in Science and Engineering*; Williams, C.J., Ed.; Kluwer Academic: Dordrecht, The Netherlands, 2003; pp. 329–349.

6. Abbas, A. Maximum Entropy Distributions with Upper and Lower Bounds. In *Bayesian Inference and Maximum Entropy Methods in Science and Engineering*; Williams, C.J., Ed.; Kluwer Academic: Dordrecht, The Netherlands, 2005; pp. 25–42.

7. Darooneh, A.H. Utility function from maximum entropy principle. *Entropy* **2006**, *8*, 18–24.

8. Dionisio, A.; Reis, A.H.; Coelho, L. Utility function estimation: The entropy approach. *Phys. A* **2008**, *387*, 3862–3867.

9. Pires, C.; Dionisio, A.; Coelho, L. Estimating utility functions using generalized maximum entropy. *J. Appl. Stat.* **2013**, *40*, 221–234.

10. Abbas, A.E. Maximum entropy utility. *Oper. Res.* **2006**, *54*, 277–290.

11. Mettke, H. Convex cubic Hermite-spline interpolation. *J. Comp. Appl. Math.* **1983**, *9*, 205–211.

12. Lai, M.J. Convex preserving scattered data interpolation using bivariate cubic splines. *J. Comp. Appl. Math.* **2000**, *119*, 249–258.

13. Visnawatahan, P.; Chand, A.K.B.; Agarwal, R.P. Preserving convexity through rational cubic spline fractal interpolation. *J. Comp. Appl. Math.* **2014**, *263*, 262–276.

14. Dybvig, P.H.; Rogers, L.C.G. Recovery of preferences from observed wealth in a single realization. *Rev. Fin. Stud.* **1997**, *10*, 151–174.

15. Cox, A.M.G.; Hobson, D.; Obłój, J. Utility theory front to back–inferring utility from agents choice's. *Int. J. Theor. Appl. Financ.* **2014**, *17*, 1450018:1–1450018:14.

16. Buchen, P.W.; Kelly, M. The maximum entropy distribution of an asset inferred from option prices. *J. Finan. Quant. Anal.* **1996**, *31*, 143–159.

17. Stutzer, M. A simple non-parametric approach to derivative security valuation. *J. Fin.* **1996**, *51*, 1633–1652.

18. Gulko, L. The entropy theory of bond option pricing. *Int. J. Theor. Appl. Fin.* **1999**, *5*, 355–383.

19. Zhou, R.; Cai, R.; Tong, G. Applications of entropy in finance: A review. *Entropy* **2013**, *15*, 4909–4931.

20. Jaynes, E.T. Information theory and statistical mechanics. *Phys. Rev.* **1957**, *106*, 620–630.

21. Gomes-Gonçalvez, E.; Gzyl, H.; Mayoral, S. Density reconstructions with errors in the data. *Entropy* **2014**, *16*, 3257–3272.

22. Navaza, J. The use of non-local constraints in maximum entropy electron density reconstruction. *Acta Crystallog.* **1986**, *42*, 212–233.

23. Dacunha-Castelle, D.; Gamboa, F. Maximum d'entropie et probleme des moments. *Ann. Inst. Henri Poincaré* **1990**, *26*, 567–596. (In French)

24. Golan, A.; Judge, G.G.; Miller, D. *Maximum Entropy Econometrics: Robust Estimation with Limited Data*; Wiley: New York, NY, USA, 1996.

25. Golan, A.; Gzyl, H. A generalized maxentropic inversion procedure for noisy data. *Appl. Math. Comp.* **2002**, *127*, 249–260.

26. Gzyl, H.; Velásquez, Y. *Linear Inverse Problems: The Maximum Entropy Connection*; World Scientific Publishers: Singapore, 2011.

27. Gzyl, H.; Mayoral, S. Determination of risk measures from market prices of risk. *Insur. Math. Econom.* **2008**, *43*, 437–443.

28. Gzyl, H.; Mayoral, S. A method for determining risk aversion from uncertain market prices of risk. *Insur. Math. Econom.* **2010**, *47*, 84–89.

29. Gzyl, H.; ter Horst, E. A relationship between the ordinary maximum entropy method and the method of maximum entropy in the mean. *Entropy* **2014**, *16*, 1123–1133.

30. Borwein, J.; Lewis, A. *Convex Analysis and Non-Linear Optimization*; CMS Books in Mathematics; Springer: New York, NY, USA, 2000.