Abstract
This paper provides a method to distribute an economic aggregate among smaller areas, using indicators. It makes use of Bayesian tools, implemented by Gibbs sampling in order to obtain the estimates. The method, with minor changes, is applied to adjust predictions provided by individual (single country) econometric models, with the overall prediction arising from an aggregate model. Authors also explore the adequacy of their method in solving problems of time series interpolation or temporal dissaggregation.

Key words: Bayesian models, Gibbs sampling, Adjust methods

1 Introduction

This paper puts forward a method to distribute an economic quantity related to an area among smaller ones. For instance, the distribution of Spanish GNP among the Autonomous Regions or the estimation of the local added value for Manufactures, knowing its amount for the whole country. The method can be used, too, in distributing an added value among several items, i.e., subsector added values.

With minor changes the method is of use when the quantity is a mean of the small areas ones instead of being their sum.

Two are the more usual topics. First, approximations to small area values are known, so the problem comes down to match these approximate values with the aggregate one. This case arises, for example, when we combine bottom-up with top-down forecasts from linked econometric models.

As a second problem, consider the situation in which we only know a growth index related to the variable's growth rate at the small area level. Such a situation emerges when we want to estimate local values, being known the related index and the country value for the magnitude.

Maybe, the method can be used for temporal interpolation or disaggregation (for example, the production of Quarterly National Accounts). Although loosely speaking this is right, changes to some extent are needed, owing to the role played by the time features of the series.

This paper poses only the first problem. Let us suppose that we know the regional values for the variable corresponding to a base year, being then know the value for the whole country ('total value' from now). We accept that the total value for the current year is known, and, also, that initial approximations to the current regional values are available. The aim is to obtain regional estimates for the current year that matches with the current total value.

---

1 From now, the small areas will be named regions, and country their sum.

The plan of the paper is as follows. In section 2 the hypothesis about priors and likelihood are posed, and the posterior distribution is derived. In section 3 a Gibbs sampler is specified and we show how it can be employed to yield the estimates. Section 4 is devoted to implementation and convergence checking and, finally, section 5 presents a numerical example from Spanish regional economic forecasting, and the performance of our estimates is checked.

2 Predictive p.d.f.’s

Let $V_0$ be the total value for the base year and $V_1$ its current value. Further, let $v_{r0}, r = 1, \ldots, P$ denote the regional values for the base year. Obviously, the base total value equals the sum of the $P$ regional values,

$$V_0 = \sum_{r=1}^{P} v_{r0}$$

Moreover, we know initial estimates, $v_{ri,ini}, r = 1, \ldots, P$, of current regional values. The aim is to obtain estimates, $v_{r1}, r = 1, \ldots, P$, matched with the totals for the current year.

We postulate two approaches to the prior distribution for the regional values: (i) we can assume that, for each region $r$, this prior is a function of $|v_{r1} - v_{ri,ini}|$, and, (ii) each prior distribution depends on the error in the estimate of growth regional rate.

In spite of the analytic simplicity of the first proposal, it leads to substantial errors for the small regions. Then, after having consider both approaches, we pronounce in favour of the second one. Consequently, denoting

$$Tv_{ri,ini} = \frac{v_{ri,ini} - v_{r0}}{v_{r0}}$$

and

$$Tv_r = \frac{v_{r1} - v_{r0}}{v_{r0}}$$

for $r = 1, \ldots, P$, we use as prior distribution a Laplace’s one (Akaike (1980) or Young (1996) use it for partially related problems), so,

$$\pi(v_{r1} \mid \lambda_r, D) \propto e^{-\lambda_r |Tv_r - Tv_{ri,ini}|}$$

being $\lambda_r$ a positive parameter and $D$ the data,

$$D = \{V_0, V_1, v_{10}, \ldots, v_{P0}, v_{1,ini}, \ldots, v_{P,ini}\}$$

Further, it is assumed that $v_{11}, \ldots, v_{P1}$ are a priori independent.

Then, the prior joint distribution for regional values results

$$\pi(v_{11}, v_{21}, \ldots, v_{P1} \mid \lambda_1, \ldots, \lambda_P, D) \propto \prod_{r=1}^{P} e^{-\lambda_r |Tv_r - Tv_{ri,ini}|}$$

We assume, too, that $\lambda_1, \ldots, \lambda_P$ are a priori independent with sparse gamma (prior) distributions (this is a common choice. See, for example, Spiegelhalter et al. (1995), Young (1996) or Zellner et al. (1991) among others)

$$\pi(\lambda_r) \propto \lambda_r e^{-\lambda_r a_r}, \quad r = 1, \ldots, P$$
with \( a_{\lambda} \), small, say \( a_{\lambda} = 0.01 \) (note that \( Var(\lambda_r) = 10^4 \) and then, prior distributions reflect our ignorance about these parameters).

The likelihood that we should employ assumes that there is a stochastic relationship between the right total growth rate and the growth rate inferred from current regional values, based on the rough nature of these later ones, so

\[
TV = TV_{\text{sum}} + \varepsilon
\]

being \( TV = \frac{V_f - V_0}{V_0} \) the total growth rate, and

\[
TV_{\text{sum}} = \frac{\sum_v v_r - V_0}{V_0}
\]

the growth rate inferred from considering the total current value as the sum of regional estimates and where the error, \( \varepsilon \), follows a Laplace (double-exponential) distribution. Thus, the likelihood is given by

\[
l(v_{11}, v_{21}, \ldots, v_{P1} | \beta, D) \propto e^{-\beta |TV - TV_{\text{sum}}|}.
\]

Finally, we consider for the positive parameter \( \beta \) a sparse gamma distribution, being independent from the others. We then take

\[
\pi(\beta) \propto \beta e^{-\beta a_{\beta}},
\]

with \( a_{\beta} = 0.01 \) as above for \( a_{\lambda} \).

Combining priors and likelihood, we obtain the posterior pdf,

\[
p(v_{11}, v_{21}, \ldots, v_{P1}, \lambda_1, \ldots, \lambda_P, \beta | D) \propto \pi(\beta) \cdot \pi(\lambda_1, \ldots, \lambda_P | \beta) \cdot \pi(v_{11}, v_{21}, \ldots, v_{P1} | \lambda_1, \ldots, \lambda_P, \beta, D) \cdot l(v_{11}, v_{21}, \ldots, v_{P1}) \propto \beta e^{-\beta a_{\beta}} \prod_{r=1}^{P} \lambda_r e^{-\lambda_r a_{\lambda_r}} \cdot e^{-\sum_r \lambda_r |Tv_r - Tv_{r,\text{ini}}|} \cdot e^{-\beta |TV - TV_{\text{sum}}|}
\]

It is known (see Ferguson (1967)) that the optimal estimate is the Bayes one, and, assuming a quadratic loss function, that estimate can be computed as the posterior expectation.

The complexity of the posterior distribution suggests the use of Gibbs sampling, a simulation algorithm pertaining to Markov Chain Monte Carlo (MCMC) models. Gamerman (1997) can be considered an up-to-day reference.

Nevertheless, the parameters \( \lambda_1, \ldots, \lambda_P, \beta \) have a volatile behavior. We have then obtained the predictive distribution for the regional values by integrating out those parameters from the last expression, with the result

\[
\int_{\beta} \int_{\lambda_1} \cdots \int_{\lambda_P} p(v_{11}, v_{21}, \ldots, v_{P1}, \lambda_1, \ldots, \lambda_P, \beta | D) d\beta d\lambda_1 \cdots d\lambda_P \propto
\]

\[
\frac{1}{(a_{\beta} + |TV - TV_{\text{sum}}|)^2} \cdot \frac{1}{P_{r=1}(a_{\lambda_r} + |Tv_r - Tv_{r,\text{ini}}|)^2}
\]

Gibbs sampling can be carried out if the conditional distributions are known (and, obviously, easy to use). Note that

\[
p(v_{h1} | v_{1}^{(h)} | D) \propto \frac{1}{(a_{\beta} + |TV - TV_{\text{sum}}|)^2} \cdot \frac{1}{(a_{\lambda_h} + |Tv_h - Tv_{h,\text{ini}}|)^2}
\]

where \( v_{1}^{(h)} \) denotes the vector \((v_{11}, v_{21}, \ldots, v_{P1})\), excluding \( v_{h1} \).
This conditional distribution can be written explicitly as a function of \( v_{h1} \), expressing \( |Tv_h - Tv_{h,ini}| \) as \( \frac{1}{v_{ho}} \cdot |v_{h1} - a| \), where \( a = v_{h,ini} \). Similarly,

\[
|TV - TV_{sum}| = \frac{1}{V_0} |v_{h1} - b|
\]

where \( b = V_1 - \sum_{r \neq h} v_{r1} \). Note that \( a \) is the initial approximation to \( v_{h1} \), and \( b \) arises from the comparison of the total value with the remaining regional values. By substitution of those expressions the conditional distribution results

\[
p(v_{h1} | v_1^{(h)}, D) \propto \frac{1}{(c + |v_{h1} - a|)^2} \cdot \frac{1}{(d + |v_{h1} - b|)^2}
\]

being

\[
a = v_{h,ini}
\]

\[
b = V_1 - \sum_{r \neq h} v_{r1}
\]

\[
c = a_\lambda \cdot v_{h0}
\]

\[
d = a_\beta \cdot V_0
\]

\[
a_\lambda, a_\beta = 0.01
\]

### 3 Simulation of regional values

In this section we set up the results supporting the simulation scheme for a regional value, say, \( v_{h1} \). To shorten the notation, we call \( x \) for \( v_{h1} \). Thus,

\[
p(x | v_1^{(h)}, D) \propto \frac{1}{(c + |x - a|)^2} \cdot \frac{1}{(d + |x - b|)^2} \equiv g(x)
\]

The simulation method is the well-known acceptation-reject one, with minor changes to speed up the iterations. Specifically we set to zero the conditional pdf for intervals with very small densities. This change diminishes the probability of rejection. Here we have the following results.

**Lemma 1** The conditional pdf for \( v_{h1} \), given \( (v_1^{(h)}, D) \) is

\[
f(x) = \frac{1}{k} \cdot g(x), \quad -\infty < x < \infty
\]

where

\[
k = \frac{2 \ln \frac{c}{M + d}}{M + d} + \frac{M + N}{c(M + d)(M + d - c)^2} + \frac{2 \ln \frac{cd}{M^2 + MN + cd}}{(M + N)^3} + \frac{M^2N + MN^2 - 2Mcd}{cd(M^2 + MN + cd)(M + N)^2} + \frac{M + N}{d(M + c)(M + c - d)^2} - \frac{2 \ln \frac{M + c}{d}}{(M + c - d)^3}
\]

being \( M = |a - b| \) and \( N = c + d \).
Proof. Integration of $g$ provides the value of $k$. ■

Lemma 2 If $a$ and $b$ are as above, it is always true that $f(a) > f(b)$. Moreover $f$ reaches its maximum\(^2\) at $a$.

Proof. Direct substitution provides

$$g(a) = \frac{1}{c^2(d + |a - b|)^2}$$

and

$$g(b) = \frac{1}{d^2(c - |a - b|)^2}$$

But

$$\frac{1}{g(a)} - \frac{1}{g(b)} = c^2(d + |a - b|)^2 - d^2(c + |a - b|)^2 = (a - b)^2(c^2 - d^2) + 2cd(c - d)|a - b|$$

and the inequality $c < d$ implies that the right member of the above expression is the sum of two negative quantities. Then $g(a) > g(b)$ and consequently $f(a) > f(b)$. The rest of the lemma can be verified in an obvious way. ■

The pdf $f$ is a decreasing function outside the interval with limits $a$ and $b$. Inside, if $a < b$, we have

$$g(x) = \frac{1}{(c + x - a)^2 \cdot (d + b - x)^2}$$

and

$$\ln(g(x)) = -2[\ln(c + x - a) + \ln(d + b - x)]$$

Differentiating this logarithm we obtain

$$\frac{d \ln(g(x))}{dx} = -2 \left[ \frac{a + b - c + d - 2x}{(c + x - a)(d + b - x)} \right]$$

This derivative is negative iff $x < \frac{a + b}{2} + \frac{d - c}{2} \left( \text{remember that } x \text{ belongs to } (a, b) \right)$.

Bearing in mind that $c < d$ and thus $\frac{d - c}{2} > 0$, the pdf is decreasing on the right of $a$, having a local minimum between $a$ and $b$ if $\frac{d - c}{2} < \frac{a + b}{2}$.

A similar reasoning can be followed for $a > b$.

Next result summarizes the above discussion.

Corollary 3 The local minimum of $f$ between $a$ and $b$ is reached, if exists, at

$$x_{\min} = \frac{a + b - (c - d)}{2}, \quad \text{if } a < b \quad \text{and } \quad a < \frac{a + b - (c - d)}{2} < b$$

or in

$$x_{\min} = \frac{a + b + (c - d)}{2}, \quad \text{if } a > b \quad \text{and } \quad a > \frac{a + b + (c - d)}{2} > b$$

Proof. See above. ■

\(^2\)Occasionally a local maximum is present at $b$. 

Proof. Integration of $g$ provides the value of $k$. ■

Lemma 2 states that $f$ reaches its maximum at $a$.

Proof. Direct substitution provides

$$g(a) = \frac{1}{c^2(d + |a - b|)^2}$$

and

$$g(b) = \frac{1}{d^2(c - |a - b|)^2}$$

But

$$\frac{1}{g(a)} - \frac{1}{g(b)} = c^2(d + |a - b|)^2 - d^2(c + |a - b|)^2 = (a - b)^2(c^2 - d^2) + 2cd(c - d)|a - b|$$

and the inequality $c < d$ implies that the right member of the above expression is the sum of two negative quantities. Then $g(a) > g(b)$ and consequently $f(a) > f(b)$. The rest of the lemma can be verified in an obvious way. ■

The pdf $f$ is a decreasing function outside the interval with limits $a$ and $b$. Inside, if $a < b$, we have

$$g(x) = \frac{1}{(c + x - a)^2 \cdot (d + b - x)^2}$$

and

$$\ln(g(x)) = -2[\ln(c + x - a) + \ln(d + b - x)]$$

Differentiating this logarithm we obtain

$$\frac{d \ln(g(x))}{dx} = -2 \left[ \frac{a + b - c + d - 2x}{(c + x - a)(d + b - x)} \right]$$

This derivative is negative iff $x < \frac{a + b}{2} + \frac{d - c}{2} \left( \text{remember that } x \text{ belongs to } (a, b) \right)$.

Bearing in mind that $c < d$ and thus $\frac{d - c}{2} > 0$, the pdf is decreasing on the right of $a$, having a local minimum between $a$ and $b$ if $\frac{d - c}{2} < \frac{a + b}{2}$.

A similar reasoning can be followed for $a > b$.

Next result summarizes the above discussion.

Corollary 3 The local minimum of $f$ between $a$ and $b$ is reached, if exists, at

$$x_{\min} = \frac{a + b - (c - d)}{2}, \quad \text{if } a < b \quad \text{and } \quad a < \frac{a + b - (c - d)}{2} < b$$

or in

$$x_{\min} = \frac{a + b + (c - d)}{2}, \quad \text{if } a > b \quad \text{and } \quad a > \frac{a + b + (c - d)}{2} > b$$

Proof. See above. ■
Lemma 4 The value of pdf at \( x_{\text{min}} \) above is

\[
f(x_{\text{min}}) = \begin{cases} 
\frac{k}{(c+x_{\text{min}}-a)^2(d+b-x_{\text{min}})^2}, & a < x_{\text{min}} < b \\
\frac{k}{(c+a-x_{\text{min}})^2(d+x_{\text{min}}-b)^2}, & b < x_{\text{min}} < a 
\end{cases}
\]

being \( k \) the constant from Lemma 1.

Proof. Substitute the value of the minimum from Corollary 3.

The appearance of pdf suggests that a direct method of simulation (such as inverse transformation, composition or convolution, among others) will be inefficient. Moreover, there is no a special method for this case. The adequacy of acceptation-rejection method comes from the fact that our pdf is bounded up. Then, it is straightforward to build up a majorizing function. Here we will describe briefly this method. Interested readers can consult either the original reference (von Neumann (1951)) or posterior versions (see, for example, Law and Kelton (1991), pp. 478-484). Without thorough intentions, applications from this method can be seen at Ahrens and Dieter (1974) (simulation of beta distributions). In this last reference authors use a piecewise linear function, as in our paper.

Acceptation-rejection method demands the setting of a mayorant function, \( t \), for \( f \), that is, \( f(x) \leq t(x) \) for all \( x \). From \( t \) we construct a pdf, \( r \), \( r \propto t \). If \( r \propto t \) exists but \( f(x_{\text{min}}) \geq p \cdot f(b) \), the majorizing \( t \) is taken as \( [\text{IALUBD}] \).

Lemma 5 Let \( p \) be a number between 0 and 1. If \( x_{\text{min}} \) exists and \( f(x_{\text{min}}) < p \cdot f(b) \), the equation \( f(x) = p \cdot f(b) \) has two roots, \( x_1 \) and \( x_2 \) into the interval:

- If \( a < x < b \),

\[
\begin{align*}
x_1 &= \frac{a + b - c + d}{2} - \sqrt{\left(\frac{a - b - c - d}{2}\right)^2 + \frac{1}{k p \cdot f(b)}} \\
x_2 &= \frac{a + b - c + d}{2} + \sqrt{\left(\frac{a - b - c - d}{2}\right)^2 + \frac{1}{k p \cdot f(b)}}
\end{align*}
\]
3 SIMULATION OF REGIONAL VALUES

If \( a < b < x < a \),

\[
x_1 = \frac{a + b + c - d}{2} - \sqrt{\left(\frac{a - b + c + d}{2}\right)^2 + \frac{1}{kp \cdot f(b)}}
\]

\[
x_2 = \frac{a + b + c - d}{2} + \sqrt{\left(\frac{a - b + c + d}{2}\right)^2 + \frac{1}{kp \cdot f(b)}}
\]

with \( x_1 < x_2 \) in both cases.

**Proof.** If \( a < b \), the equation

\[
f(x) = \frac{1}{k \cdot (c + x - a)^2(d + b - x)^2} = p \cdot f(b)
\]

can be written as

\[
(c + x - a)^2(d + b - x)^2 - \frac{1}{kp \cdot f(b)} = 0
\]

By solving this last equation we obtain four roots, namely \( x_1, x_2 \) as above, and the other two are outside \((a, b)\). Inequality \( x_1 < x_2 \) is obvious.

The proof for \( b < a \) is analogous. ■

The following lemmas set up the cutoff points at the both sides of the interval.

**Lemma 6** Let \( p \) a number between 0 and 1. A root of \( f(x) = p \cdot f(b) \) stays on the left of \( \min(a, b) \) and results

\[
x_{izq} = a + b + c + d - \sqrt{\left(\frac{a - b + c - d}{2}\right)^2 + \frac{1}{p \cdot k \cdot f(b)}}
\]

**Proof.** If \( x < \min(a, b) \) the equation

\[
f(x) = \frac{1}{k \cdot (c - x - a)^2(d + b - x)^2} = p \cdot f(b)
\]

can be written as

\[
(x - (a + c))^2(x - (b + d))^2 - \frac{1}{kp \cdot f(b)} = 0
\]

and the result follows. ■

**Lemma 7** Let \( p \) a number between 0 and 1. A root of \( f(x) = p \cdot f(b) \) stays on the right of \( \max(a, b) \) and is given by

\[
x_{der} = a + b - c - d + \sqrt{\left(\frac{a - b - c + d}{2}\right)^2 + \frac{1}{kp \cdot f(b)}}
\]

**Proof.** Analogous to the above lemma. ■
4 Implementation and convergence checking

As it is known, there is a wide controversy into the field of MCMC models related to the initial values for the chain, quasi-independence or testing for convergence. The above reference from Gamerman (Gamerman (1997)) assembles the more relevant results and a brief discussion among procedures.

Generally speaking, there is a trade-off between efficiency and speed, depending on the size of the problem. Here, the high number of variables leads to simple and self-acting diagnosis methods. Specifically:

- **Initial values and number of chains.** Gelman and Rubin (1992) have shown the advantages of multiple chains in avoiding the concentration of sample values around any local mode, and in making easier the convergence diagnostics. Our choice points at a single chain with 3000 iterations, due to the high number of variables and the certainty of having satisfactory initial values (not being far from correct values). Specifically, the initial value \(v_{r}^{0}\) for a region, \(r\), is settled as the arithmetic mean of two approximations. The first one
  \[v_{r}^{1} = v_{r0} \cdot \frac{V_{1}}{V_{0}},\]
emerges after applying the total growth rate to regional base values. The second one, \(v_{r}^{2}\), is the initial value \(v_{r,ini}\). Finally,
  \[(0)v_{r} = \frac{v_{r}^{1} + v_{r}^{2}}{2}.\]

- **Convergence diagnosis.** Here a lot of work has also been made, and the book of Gamerman (1997) is a good reference. Again our choose reflects the high number of variables implied and the need of an automatic procedure. Our proposal follows Geweke (1992). Given a sequence \(\{v_{h1}(j)\}_{j=1}^{n}\), let \(n_{e} (n_{l})\) denote respectively a number of early (late) values. Then, as \(n \rightarrow \infty\)
  \[(v_{h1} - \bar{v}_{h1})/\sqrt{\hat{V}_{e} + \hat{V}_{l}} \sim N(0, 1)\]
being \((n_{e} + n_{l})/n < 1\), \(n_{e}/n\) and \(n_{l}/n\) fixed ratios, \(v_{h1} = \bar{v}_{h1}\) the sample mean of early (late) values, and \(\hat{V}_{e}\) (\(\hat{V}_{l}\)) estimates of the variances.

- **Confidence intervals.** Even though a confidence interval can be computed from the sample values, a more efficient method implies the use of conditional distribution, computing previously an estimator of the marginal posterior pdf Rao-Blackwellized (see, for example, Gelfand and Smith (1990)). Specifically
  \[f_{v_{h1}}(x | D) = \frac{1}{n} \sum_{j=1}^{n} f(x | (j)v_{1}^{(h)}; D)\]
where \(j\) denotes the iteration number, and we use the last \(n\) iterations (\(n = 200\) in our application).
This pdf is used to estimate the variance and the percentiles at 5% and 95%.

- **Implementation.** The procedure has been implemented by using TSP, v.4.2b.

5 An estimation of Spanish regional output

This section is devoted to apply the above method in obtaining estimates of regional added values for the 18 Spanish regions. The Spanish National Statistics Institute (INE) provides the total added value for year 1999 (see National
5 AN ESTIMATION OF SPANISH REGIONAL OUTPUT

Accounts, i.e. Contabilidad Nacional de España (2000)). Moreover the HISPALINK Group\(^3\) estimates the regional added values, but they are unmatched with the total one.

Our congruence method applies both to the overall regional values and to the disaggregate regional values for six sectors: Agriculture, Energy, Manuamufacturing goods, Building, Market Services and Government Services.

Obviously, the initial estimates use a forecasted total added value (from L.R. Klein Institute) as input. Then, the error in the sum of regional initial estimates is of (minor importance) little extent (Table 2 at the end of this paper shows these errors that range from 0.01\% (Market Services) to 2.95\% (Energy)).

We also know the regional sectorial output for year 1998, provided for HISPADAT (HISPADAT is the databank of the HISPALINK Group, see HISPALINK (1999)). These values are considered as given, and match with the Spanish output for that year.

Table 2 at the end of the paper shows the estimated output for each region and sector. They are computed as the ergodic means after performing the Gibbs sampler. Also, that table provides estimates of standard deviations, allowing approximately normal confidence intervals to be built. These deviations are computed from the estimated marginal pdf of each variable, and not directly from the chain.

Gamerman (1997) suggests computing posterior means from the estimated marginal posterior pdf, as for the standard deviation. However, the approximations used for that pdf estimation produce a worse matching with the totals. Table 1 below shows this fact, being the errors from these last estimations very similar to the initial ones.

Table 1. Errors arising from estimates

Comparison between the sum of regional values and the total

<table>
<thead>
<tr>
<th></th>
<th>Initial values</th>
<th>Ergodic mean estimates</th>
<th>Estimates from marginal predictive pdf</th>
</tr>
</thead>
<tbody>
<tr>
<td>All sectors</td>
<td>0.54</td>
<td>0.13</td>
<td>0.54</td>
</tr>
<tr>
<td>Agriculture</td>
<td>0.80</td>
<td>0.22</td>
<td>0.72</td>
</tr>
<tr>
<td>Energy</td>
<td>2.95</td>
<td>1.23</td>
<td>2.74</td>
</tr>
<tr>
<td>Manufacturing goods</td>
<td>1.01</td>
<td>0.30</td>
<td>0.90</td>
</tr>
<tr>
<td>Building</td>
<td>1.53</td>
<td>0.52</td>
<td>1.37</td>
</tr>
<tr>
<td>Market Services</td>
<td>0.01</td>
<td>-0.04</td>
<td>0.01</td>
</tr>
<tr>
<td>Government Services</td>
<td>0.38</td>
<td>0.07</td>
<td>0.36</td>
</tr>
</tbody>
</table>

Table 1 shows also that the ergodic means are clearly better than the others for all sectors, with the exception of the Market Services.

Furthermore, we provide a 90\% confidence interval, arising through the 5\% and 95\% percentiles. Here also, the quality of intervals suffers from the approximation used in estimating the marginal pdf. So, if

\[
\min v_{h1} = \min_{j=1,\ldots,n} (j) v_{h1} \quad \text{and} \quad \max v_{h1} = \max_{j=1,\ldots,n} (j) v_{h1}
\]

the interval \((\min v_{h1}, \max v_{h1})\) is partitioned into 200 subintervals and for each, an estimate of the value of the marginal pdf is computed as the mean of the conditional densities for simulated values of the other variables. Bearing in mind

\(^3\)HISPALINK is a project from fifteen universities that assembles 15 groups devoted to the forecasting of the economic behaviour of Spanish regions. Their forecast’s link is of bottom-up type. Then, after individual forecasting, a second step is devoted to match the regional estimates with the total values. Our paper tackles this second phase.
that conditional densities are highly leptokurtics, the accuracy of the approximation will suffer around the modal values.

Table 3 at the end of the paper shows the differences between the initial growth rates and the ones computed from our estimates. The adjustments point to the same direction owing to the initial underestimation of the totals.

6 Summary and concluding remarks

In this paper, we have provided a method in order to estimate economic quantities for small areas, being known both initial estimates and the totals for the aggregated area. We obtain a Bayes estimator from almost non-informative prior distributions and double-exponential likelihood. The estimation is carried on via Gibbs sampling from the conditional distributions.

Moreover, we compute precision measures for the estimators, using estimates of the marginal posteriors.

The method applies to match initial estimates of the regional output for Spanish Autonomous Regions with the total output, substantially improving the initial results.

Note that the availability of initial values is not necessary. It is enough for us to know the regional growth rates for an appropriate coincidental index.

Even though the method can be applied sequentially along several years, the procedure is not optimal. Actually we are working with such an extension. Other improvement lies on superimposing restrictions on the values, so that the yearly growth rate stands on the grounds suggested by experts. This improvement in not difficult.

7 Bibliography


8 Annexe. Tables and figures

Figure 1. Conditional pdf and majorizing function
<table>
<thead>
<tr>
<th>Region</th>
<th>Total</th>
<th>Agriculture</th>
<th>Energy</th>
<th>Manufacturing</th>
<th>Goods</th>
<th>Building</th>
<th>Market Services</th>
<th>Government Services</th>
</tr>
</thead>
<tbody>
<tr>
<td>Andalucía</td>
<td>110 910.0</td>
<td>389 040.0</td>
<td>123 959.0</td>
<td>108 949.0</td>
<td>123 590.0</td>
<td>937 890.0</td>
<td>179 590.0</td>
<td>1 075 900.0</td>
</tr>
<tr>
<td>Extremadura</td>
<td>8 820.0</td>
<td>2 980.0</td>
<td>1 840.0</td>
<td>1 320.0</td>
<td>1 780.0</td>
<td>15 280.0</td>
<td>2 820.0</td>
<td>32 820.0</td>
</tr>
<tr>
<td>Com. de Madrid</td>
<td>124 010.0</td>
<td>46 810.0</td>
<td>34 620.0</td>
<td>24 400.0</td>
<td>28 200.0</td>
<td>113 210.0</td>
<td>20 610.0</td>
<td>136 820.0</td>
</tr>
<tr>
<td>Galicia</td>
<td>100 410.0</td>
<td>36 810.0</td>
<td>26 420.0</td>
<td>18 200.0</td>
<td>22 800.0</td>
<td>95 810.0</td>
<td>17 610.0</td>
<td>113 420.0</td>
</tr>
<tr>
<td>Region de Murcia</td>
<td>183 970.0</td>
<td>61 840.0</td>
<td>42 350.0</td>
<td>32 140.0</td>
<td>36 740.0</td>
<td>168 350.0</td>
<td>30 740.0</td>
<td>199 050.0</td>
</tr>
<tr>
<td>País Vasco</td>
<td>86 650.0</td>
<td>30 940.0</td>
<td>22 430.0</td>
<td>16 220.0</td>
<td>19 820.0</td>
<td>69 050.0</td>
<td>12 850.0</td>
<td>81 900.0</td>
</tr>
<tr>
<td>La Rioja</td>
<td>174 310.0</td>
<td>57 430.0</td>
<td>38 240.0</td>
<td>28 040.0</td>
<td>32 640.0</td>
<td>158 710.0</td>
<td>27 310.0</td>
<td>186 020.0</td>
</tr>
<tr>
<td>País Vasco</td>
<td>86 650.0</td>
<td>30 940.0</td>
<td>22 430.0</td>
<td>16 220.0</td>
<td>19 820.0</td>
<td>69 050.0</td>
<td>12 850.0</td>
<td>81 900.0</td>
</tr>
</tbody>
</table>

### Table 2: 2. Market estimates and dispersion measures

<table>
<thead>
<tr>
<th>Region</th>
<th>σ0.05</th>
<th>σ0.95</th>
<th>σ0.05</th>
<th>σ0.95</th>
<th>σ0.05</th>
<th>σ0.95</th>
<th>σ0.05</th>
<th>σ0.95</th>
</tr>
</thead>
<tbody>
<tr>
<td>Andalucía</td>
<td>100 410.0</td>
<td>36 810.0</td>
<td>26 420.0</td>
<td>18 200.0</td>
<td>22 800.0</td>
<td>95 810.0</td>
<td>17 610.0</td>
<td>113 420.0</td>
</tr>
<tr>
<td>Extremadura</td>
<td>8 820.0</td>
<td>2 980.0</td>
<td>1 840.0</td>
<td>1 320.0</td>
<td>1 780.0</td>
<td>15 280.0</td>
<td>2 820.0</td>
<td>32 820.0</td>
</tr>
<tr>
<td>Com. de Madrid</td>
<td>124 010.0</td>
<td>46 810.0</td>
<td>34 620.0</td>
<td>24 400.0</td>
<td>28 200.0</td>
<td>113 210.0</td>
<td>20 610.0</td>
<td>136 820.0</td>
</tr>
<tr>
<td>Galicia</td>
<td>100 410.0</td>
<td>36 810.0</td>
<td>26 420.0</td>
<td>18 200.0</td>
<td>22 800.0</td>
<td>95 810.0</td>
<td>17 610.0</td>
<td>113 420.0</td>
</tr>
<tr>
<td>Region de Murcia</td>
<td>183 970.0</td>
<td>61 840.0</td>
<td>42 350.0</td>
<td>32 140.0</td>
<td>36 740.0</td>
<td>168 350.0</td>
<td>30 740.0</td>
<td>199 050.0</td>
</tr>
<tr>
<td>País Vasco</td>
<td>86 650.0</td>
<td>30 940.0</td>
<td>22 430.0</td>
<td>16 220.0</td>
<td>19 820.0</td>
<td>69 050.0</td>
<td>12 850.0</td>
<td>81 900.0</td>
</tr>
<tr>
<td>La Rioja</td>
<td>174 310.0</td>
<td>57 430.0</td>
<td>38 240.0</td>
<td>28 040.0</td>
<td>32 640.0</td>
<td>158 710.0</td>
<td>27 310.0</td>
<td>186 020.0</td>
</tr>
</tbody>
</table>

### 8 ANNEXE. TABLES AND FIGURES
Table 3. Real growth rate for regional output
(Estimated and initial values)

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Andalucía</td>
<td>3.3</td>
<td>3.6</td>
<td>-4.9</td>
<td>-4.5</td>
<td>-1.5</td>
<td>0.4</td>
<td>0.6</td>
<td>1.5</td>
</tr>
<tr>
<td>Aragón</td>
<td>2.2</td>
<td>2.8</td>
<td>-7.1</td>
<td>-6.5</td>
<td>-2.1</td>
<td>0.1</td>
<td>1.6</td>
<td>2.7</td>
</tr>
<tr>
<td>Asturias</td>
<td>2.9</td>
<td>1.6</td>
<td>2.4</td>
<td>3.2</td>
<td>8.8</td>
<td>3.0</td>
<td>1.8</td>
<td>2.7</td>
</tr>
<tr>
<td>Baleares</td>
<td>5.5</td>
<td>6.1</td>
<td>-8.9</td>
<td>-8.3</td>
<td>7.0</td>
<td>8.8</td>
<td>2.9</td>
<td>4.0</td>
</tr>
<tr>
<td>Canarias</td>
<td>4.3</td>
<td>4.8</td>
<td>-3.8</td>
<td>-3.1</td>
<td>2.5</td>
<td>4.3</td>
<td>3.0</td>
<td>4.0</td>
</tr>
<tr>
<td>Cantabria</td>
<td>2.9</td>
<td>3.3</td>
<td>3.7</td>
<td>4.5</td>
<td>1.1</td>
<td>3.1</td>
<td>1.6</td>
<td>2.7</td>
</tr>
<tr>
<td>Castilla y León</td>
<td>3.1</td>
<td>3.6</td>
<td>4.0</td>
<td>4.5</td>
<td>-0.9</td>
<td>0.7</td>
<td>2.9</td>
<td>3.6</td>
</tr>
<tr>
<td>Castilla La Mancha</td>
<td>3.2</td>
<td>3.8</td>
<td>2.1</td>
<td>2.8</td>
<td>2.9</td>
<td>4.6</td>
<td>2.7</td>
<td>3.5</td>
</tr>
<tr>
<td>Cataluña</td>
<td>3.6</td>
<td>3.8</td>
<td>-0.6</td>
<td>0.1</td>
<td>0.2</td>
<td>1.9</td>
<td>3.2</td>
<td>3.6</td>
</tr>
<tr>
<td>Com. Valenciana</td>
<td>3.0</td>
<td>3.5</td>
<td>-5.8</td>
<td>-5.2</td>
<td>1.6</td>
<td>3.5</td>
<td>3.6</td>
<td>4.4</td>
</tr>
<tr>
<td>Extremadura</td>
<td>2.7</td>
<td>3.2</td>
<td>-2.4</td>
<td>-1.7</td>
<td>-0.1</td>
<td>1.5</td>
<td>2.1</td>
<td>3.2</td>
</tr>
<tr>
<td>Galicia</td>
<td>3.4</td>
<td>4.1</td>
<td>0.9</td>
<td>1.5</td>
<td>2.4</td>
<td>4.2</td>
<td>4.2</td>
<td>5.1</td>
</tr>
<tr>
<td>Com. de Madrid</td>
<td>3.5</td>
<td>3.9</td>
<td>-9.0</td>
<td>-7.9</td>
<td>0.3</td>
<td>2.2</td>
<td>3.2</td>
<td>3.9</td>
</tr>
<tr>
<td>Región de Murcia</td>
<td>3.2</td>
<td>3.7</td>
<td>-1.8</td>
<td>-1.0</td>
<td>2.0</td>
<td>3.7</td>
<td>1.9</td>
<td>2.9</td>
</tr>
<tr>
<td>Com. F. Navarra</td>
<td>2.9</td>
<td>3.6</td>
<td>-1.2</td>
<td>-0.5</td>
<td>-2.1</td>
<td>-0.2</td>
<td>3.6</td>
<td>4.4</td>
</tr>
<tr>
<td>País Vasco</td>
<td>2.5</td>
<td>2.9</td>
<td>-6.3</td>
<td>-5.7</td>
<td>1.4</td>
<td>3.1</td>
<td>0.7</td>
<td>1.5</td>
</tr>
<tr>
<td>La Rioja</td>
<td>3.0</td>
<td>3.4</td>
<td>2.2</td>
<td>3.0</td>
<td>-1.8</td>
<td>-0.1</td>
<td>2.3</td>
<td>3.3</td>
</tr>
<tr>
<td>Ceuta y Melilla</td>
<td>3.0</td>
<td>3.2</td>
<td>-3.9</td>
<td>-1.3</td>
<td>0.8</td>
<td>2.6</td>
<td>0.0</td>
<td>3.2</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Andalucía</td>
<td>14.9</td>
<td>16.0</td>
<td>5.1</td>
<td>5.1</td>
<td>0.2</td>
<td>0.4</td>
</tr>
<tr>
<td>Aragón</td>
<td>7.3</td>
<td>8.6</td>
<td>3.5</td>
<td>3.6</td>
<td>1.8</td>
<td>2.1</td>
</tr>
<tr>
<td>Asturias</td>
<td>7.8</td>
<td>8.9</td>
<td>2.0</td>
<td>2.4</td>
<td>0.5</td>
<td>0.9</td>
</tr>
<tr>
<td>Baleares</td>
<td>11.6</td>
<td>12.8</td>
<td>6.7</td>
<td>6.9</td>
<td>0.7</td>
<td>1.0</td>
</tr>
<tr>
<td>Canarias</td>
<td>10.0</td>
<td>11.2</td>
<td>5.3</td>
<td>5.4</td>
<td>1.5</td>
<td>1.9</td>
</tr>
<tr>
<td>Cantabria</td>
<td>5.9</td>
<td>7.2</td>
<td>3.0</td>
<td>3.5</td>
<td>0.9</td>
<td>1.5</td>
</tr>
<tr>
<td>Castilla y León</td>
<td>5.8</td>
<td>6.9</td>
<td>3.9</td>
<td>3.9</td>
<td>1.6</td>
<td>2.0</td>
</tr>
<tr>
<td>Castilla La Mancha</td>
<td>4.8</td>
<td>6.0</td>
<td>4.7</td>
<td>4.7</td>
<td>0.6</td>
<td>1.0</td>
</tr>
<tr>
<td>Cataluña</td>
<td>6.8</td>
<td>7.8</td>
<td>4.0</td>
<td>4.0</td>
<td>1.3</td>
<td>1.6</td>
</tr>
<tr>
<td>Com. Valenciana</td>
<td>6.5</td>
<td>7.7</td>
<td>3.5</td>
<td>3.6</td>
<td>1.4</td>
<td>1.8</td>
</tr>
<tr>
<td>Extremadura</td>
<td>7.1</td>
<td>8.3</td>
<td>3.7</td>
<td>4.2</td>
<td>1.5</td>
<td>1.9</td>
</tr>
<tr>
<td>Galicia</td>
<td>10.2</td>
<td>11.2</td>
<td>3.2</td>
<td>3.3</td>
<td>2.4</td>
<td>2.8</td>
</tr>
<tr>
<td>Com. de Madrid</td>
<td>9.5</td>
<td>10.5</td>
<td>4.0</td>
<td>3.9</td>
<td>1.0</td>
<td>1.2</td>
</tr>
<tr>
<td>Región de Murcia</td>
<td>13.3</td>
<td>14.6</td>
<td>3.4</td>
<td>3.7</td>
<td>1.5</td>
<td>2.0</td>
</tr>
<tr>
<td>Com. F. Navarra</td>
<td>2.8</td>
<td>4.1</td>
<td>3.8</td>
<td>4.0</td>
<td>1.4</td>
<td>1.9</td>
</tr>
<tr>
<td>País Vasco</td>
<td>2.0</td>
<td>3.3</td>
<td>4.8</td>
<td>4.9</td>
<td>0.9</td>
<td>1.4</td>
</tr>
<tr>
<td>La Rioja</td>
<td>5.1</td>
<td>6.4</td>
<td>3.9</td>
<td>3.9</td>
<td>1.0</td>
<td>1.8</td>
</tr>
<tr>
<td>Ceuta y Melilla</td>
<td>8.4</td>
<td>9.7</td>
<td>3.8</td>
<td>4.2</td>
<td>1.0</td>
<td>1.4</td>
</tr>
</tbody>
</table>