Technical Report: The functional singular value decomposition for bivariate stochastic processes

Daniel Gervini
Department of Mathematical Sciences
University of Wisconsin–Milwaukee

January 26, 2009

1 Proofs of theorems

Given the stochastic process mean $\mu \in L^2(\mathcal{S} \times \mathcal{T})$, define the operator $\mathfrak{M} : L^2(\mathcal{T}) \to L^2(\mathcal{S})$ as

$$(\mathfrak{M}f)(s) = \int_{\mathcal{T}} \mu(s,t) f(t) \mathrm{d}t.$$

The adjoint of $\mathfrak{M}$ is the operator $\mathfrak{M}^* : L^2(\mathcal{S}) \to L^2(\mathcal{T})$ given by

$$(\mathfrak{M}^* g)(t) = \int_{\mathcal{S}} \mu(s,t) g(s) \mathrm{d}s.$$

Let $\mathfrak{K}_1 = \mathfrak{M}\mathfrak{M}^*$ and $\mathfrak{K}_2 = \mathfrak{M}^*\mathfrak{M}$, which are self-adjoint operators, with $\mathfrak{K}_1 : L^2(\mathcal{S}) \to L^2(\mathcal{S})$ and $\mathfrak{K}_2 : L^2(\mathcal{T}) \to L^2(\mathcal{T})$. These are integral operators with kernels $k_1(s_1, s_2) = \int \mu(s_1, t)\mu(s_2, t)\mathrm{d}t$ and $k_2(t_1, t_2) = \int \mu(s, t_1)\mu(s, t_2)\mathrm{d}s$, respectively.

The next proof will use functional analysis concepts and results that can be found, for instance, in Gohberg et al. (2003). Remember that for $f \in \mathcal{H}_1$ and $g \in \mathcal{H}_2$, the tensor-product operator $g \otimes f : \mathcal{H}_1 \to \mathcal{H}_2$ is defined as $(g \otimes f)(h) = \langle f, h \rangle g$.

1.1 Proof of Theorem 1

Since $\mathfrak{K}_2$ is a self-adjoint integral (hence compact) operator, the well-known spectral decomposition implies that $\mathfrak{K}_2 = \sum \lambda_k \psi_k \otimes \psi_k$, where $\lambda_k > 0$ and $\{\psi_k\}$ is an orthonormal system of eigenfunctions of $\mathfrak{K}_2$, which can be completed to a basis of $L^2(\mathcal{T})$ by adding an orthonormal basis of $\ker(\mathfrak{K}_2)$, say $\{\tilde{\psi}_k\}$ (Gohberg et al., 2003, p. 180). This proves (2) of Theorem 1.

Note that $\ker(\mathfrak{K}_2) = \ker(\mathfrak{M})$: clearly $\ker(\mathfrak{M}) \subseteq \ker(\mathfrak{K}_2)$ because $\mathfrak{K}_2 = \mathfrak{M}^*\mathfrak{M}$; but for any $f \in \ker(\mathfrak{K}_2)$, $0 = \langle f, \mathfrak{K}_2 f \rangle = \|\mathfrak{M}f\|^2$, which implies $f \in \ker(\mathfrak{M})$ and then $\ker(\mathfrak{K}_2) \subseteq \ker(\mathfrak{M})$.

Now define

$$\phi_k = \frac{1}{\lambda_k^{1/2}} \mathfrak{M}\psi_k.$$

The ϕ_ks are orthonormal in $L^2(\mathcal{S})$, since

$$\begin{aligned} \langle \phi_j, \phi_k \rangle &= \frac{1}{\lambda_j^{1/2}\lambda_k^{1/2}} \langle \mathfrak{M}\psi_j, \mathfrak{M}\psi_k \rangle \\ &= \frac{1}{\lambda_j^{1/2}\lambda_k^{1/2}} \langle \psi_j, \mathfrak{K}_2\psi_k \rangle \\ &= \frac{1}{\lambda_j^{1/2}\lambda_k^{1/2}} \lambda_k \delta_{jk}. \end{aligned}$$

To prove (3) of Theorem 1, define the operator $\mathfrak{L} = \sum \lambda_k^{1/2} \phi_k \otimes \psi_k$. This operator is well defined, since for any $f \in L^2(\mathcal{T})$, we have $\mathfrak{L}f = \sum \lambda_k^{1/2} \langle \psi_k, f \rangle \phi_k$ and

$$\|\mathfrak{L}f\|^2 = \sum \lambda_k |\langle \psi_k, f \rangle|^2 \leq \|f\|^2 \sum \lambda_k < \infty.$$

Direct calculation shows that $\mathfrak{L}\psi_k = \mathfrak{M}\psi_k$, and $\mathfrak{L}\tilde{\psi}_k = \mathfrak{M}\tilde{\psi}_k = 0$ because $\ker(\mathfrak{K}_2) = \ker(\mathfrak{M})$. Since $\{\psi_k\} \cup \{\tilde{\psi}_k\}$ is a basis of $L^2(\mathcal{T})$, it follows that $\mathfrak{L} = \mathfrak{M}$, which is (3) of Theorem 1 in different words.

The identity (1) of Theorem 1 follows from (3), since $\mathfrak{K}_1 = \mathfrak{M}\mathfrak{M}^*$. In particular, this shows that the positive eigenvalues of $\mathfrak{K}_1$ are the same as those of $\mathfrak{K}_2$, and the ϕ_ks can be taken as the corresponding eigenfunctions.

If the mean function $\mu(s,t)$ is continuous, by Mercer's theorem (Gohberg et al., 2003) we have that the ψ_ks are continuous and k_2 satisfies (2) in Theorem 1 in a pointwise manner, with the series converging absolutely and uniformly.

The ϕ_ks are continuous by definition when μ is continuous. To prove that the identity (1) in Theorem 1 holds pointwise and that the series converges absolutely and uniformly, we essentially have to mimic the proof of Mercer's theorem (Gohberg et al., 2003, p. 198). Since

$$\sum_{k=1}^{n} |\lambda_k \phi_k(s_1)\phi_k(s_2)| \leq \Big\{ \sum_{k=1}^{n} \lambda_k \phi_k^2(s_1) \Big\}^{1/2} \Big\{ \sum_{k=1}^{n} \lambda_k \phi_k^2(s_2) \Big\}^{1/2},$$

the uniform bound

$$\sum_{k=1}^{n} \lambda_k \phi_k^2(s) \leq \max_{x \in \mathcal{S}} k_1(x,x) \tag{1}$$

(derived below) implies that

$$c(s_1,s_2) = \sum \lambda_k \phi_k(s_1)\phi_k(s_2)$$

is well defined and continuous in $\mathcal{S} \times \mathcal{S}$, because the series converges absolutely and uniformly in both variables. If $\mathfrak{C} : L^2(\mathcal{S}) \to L^2(\mathcal{S})$ is the integral operator with kernel $c(s_1,s_2)$, direct calculation shows that $\mathfrak{K}_1 f = \mathfrak{C} f$ for all $f \in L^2(\mathcal{S})$. Then $k_1(s_1,s_2) = c(s_1,s_2)$ almost everywhere, but since k_1 and c are continuous, the equality must hold everywhere in $\mathcal{S} \times \mathcal{S}$ and identity (1) in Theorem 1 follows.

To derive (1), define

$$r_n(s_1,s_2) = k_1(s_1,s_2) - \sum_{k=1}^{n} \lambda_k \phi_k(s_1)\phi_k(s_2),$$

which is continuous, symmetric, and satisfies

$$\iint r_n(s_1,s_2) f(s_1) f(s_2)\, ds_1\, ds_2 = \sum_{k>n} \lambda_k \left|\langle \phi_k, f\rangle\right|^2 \geq 0$$

for any $f \in L^2(\mathcal{S})$. Then, by Lemma 2.1 on p. 196 of Gohberg et al. (2003), $r_n(x,x) \geq 0$ for all $x \in \mathcal{S}$ and (1) follows.

Finally, to show that expression (3) in Theorem 1 holds pointwise when the series on the right-hand side converges absolutely and uniformly, note that both sides of expression (3) define the same operator from $L^2(\mathcal{T})$ to $L^2(\mathcal{S})$, so the identity must hold almost everywhere, and by continuity, it must actually hold everywhere.■

Remark. As by-products of the proof of Theorem 1 we obtain the identities

$$\phi_k(s) = \frac{1}{\lambda_k^{1/2}}(\mathfrak{M}\psi_k)(s) = \frac{1}{\lambda_k^{1/2}}\int \mu(s,t)\psi_k(t)\mathrm{d}t,$$

and

$$\psi_k(t) = \frac{1}{\lambda_k^{1/2}}(\mathfrak{M}^*\phi_k)(t) = \frac{1}{\lambda_k^{1/2}}\int \mu(s,t)\phi_k(s)\mathrm{d}s.$$

1.2 Proof of Theorem 2

Since $\{f_k\}$ and $\{g_k\}$ are orthonormal in their respective spaces,

$$\|\mu - h\|^2 = \|\mu\|^2 - 2\sum_{k=1}^{p} a_k\langle g_k, \mathfrak{M}f_k\rangle + \sum_{k=1}^{p} a_k^2,$$

which is minimized by $a_k = \langle g_k, \mathfrak{M}f_k\rangle,\ k = 1,\ldots,p$. Then, minimizing $\|\mu - h\|^2$ is equivalent to maximizing $\sum_{k=1}^{p} |\langle g_k, \mathfrak{M}f_k\rangle|^2$. By Cauchy-Schwartz inequality,

$$\begin{aligned}\sum_{k=1}^{p} |\langle g_k, \mathfrak{M}f_k\rangle|^2 &\leq \sum_{k=1}^{p} \|g_k\|^2 \|\mathfrak{M}f_k\|^2 \\ &= \sum_{k=1}^{p} |\langle \mathfrak{M}f_k, \mathfrak{M}f_k\rangle|^2 \\ &= \sum_{k=1}^{p} |\langle f_k, \mathfrak{K}_2 f_k\rangle|^2 . \qquad (2)\end{aligned}$$

It is well known (or see Gohberg et al., 2003, sec. 4.9) that (2) is maximized by the leading p eigenfunctions of $\mathfrak{K}_2$, and the maximum value is $\sum_{k=1}^{p} \lambda_k$. Therefore,

$$\sum_{k=1}^{p} |\langle g_k, \mathfrak{M}f_k\rangle|^2 \leq \sum_{k=1}^{p} \lambda_k$$

and equality holds for $f_k = \psi_k$ and $g_k = \phi_k$, which completes the proof. ∎

1.3 Proof of Theorem 3

Let $z_{ijk} = x_{ijk} - \mu(s_j, t_k)$, and define $\mathbf{M}_0 = [\mu(s_j, t_k)]_{(j,k)}$, $\mathbf{X}_i = [x_{ijk}]_{(j,k)}$ and $\mathbf{Z}_i = [z_{ijk}]_{(j,k)}$. Since $\hat{\mathbf{\Omega}} = \mathbf{B}^\top\mathbf{V}\mathbf{K}_1\mathbf{V}\mathbf{B}$ and $\mathbf{K}_1 = \bar{\mathbf{X}}\mathbf{U}\bar{\mathbf{X}}^\top$, we can write

$$\begin{aligned}\hat{\Omega}_{hh'} &= \beta_h(\mathbf{s})^\top\mathbf{V}\bar{\mathbf{X}}\mathbf{U}\bar{\mathbf{X}}^\top\mathbf{V}\beta_{h'}(\mathbf{s}) \\ &= \beta_h(\mathbf{s})^\top\mathbf{V}\mathbf{M}_0\mathbf{U}\mathbf{M}_0^\top\mathbf{V}\beta_{h'}(\mathbf{s}) \qquad (3)\\ &\quad +2\beta_h(\mathbf{s})^\top\mathbf{V}\bar{\mathbf{Z}}\mathbf{U}\mathbf{M}_0^\top\mathbf{V}\beta_{h'}(\mathbf{s}) \qquad (4)\\ &\quad +\beta_h(\mathbf{s})^\top\mathbf{V}\bar{\mathbf{Z}}\mathbf{U}\bar{\mathbf{Z}}^\top\mathbf{V}\beta_{h'}(\mathbf{s}). \qquad (5)\end{aligned}$$

We will show that (3) goes to $\Omega_{hh'}$ as m and r go to infinity, and that (4) and (5) go to zero in probability as n goes to infinity, uniformly in m and r.

Since

$$\beta_h(\mathbf{s})^\top \mathbf{V}\bar{\mathbf{X}}\mathbf{U}\bar{\mathbf{X}}^\top \mathbf{V}\beta_{h'}(\mathbf{s}) =$$
$$\sum_{j=1}^m \sum_{j'=1}^m \beta_h(s_j) v_j \{\sum_{k=1}^r u_k \mu(s_j, t_k)\mu(s_{j'}, t_k)\} v_{j'} \beta_{h'}(s_{j'}),$$

it is clear that (3) goes to $\Omega_{hh'}$ as m and r go to infinity, because both $\max v_j$ and $\max u_k$ go to zero as m and r go to infinity.

With respect to (4), note that we can write it as $2\bar{y}$, with

$$y_i = \beta_h(\mathbf{s})^\top \mathbf{V}\mathbf{Z}_i \mathbf{U}\mathbf{M}_0^\top \mathbf{V}\beta_{h'}(\mathbf{s}).$$

The y_is are i.i.d. with $\mathrm{E}(y_i) = 0$ and

$$\mathrm{V}(y_i) = \mathrm{V}\{\sum_j \sum_k \beta_h(s_j) v_j z_{ijk} u_k a_{kh'}\},$$

with $a_{kh'} = \sum_{j'} \mu(s_{j'}, t_k) v_{j'} \beta_{h'}(s_{j'})$. Note that $a_{kh'} \to \int \mu(s, t_k)\beta_{h'}(s)\mathrm{d}s$ as $m \to \infty$; let us call this integral $\alpha_{h'}(t_k)$. Since

$$\mathrm{Cov}(z_{ijk}, z_{ij'k'}) = \rho\{(s_j, t_k), (s_{j'}, t_{k'})\} + \sigma^2 \delta_{jj'}\delta_{kk'},$$

we have

$$\begin{aligned} \mathrm{V}(y_i) &= \sum_j \sum_k \sum_{j'} \sum_{k'} \beta_h(s_j) v_j u_k a_{kh'} \beta_h(s_{j'}) v_{j'} u_{k'} a_{k'h'} \rho\{(s_j, t_k), (s_{j'}, t_{k'})\} \\ &\quad + \sigma^2 \sum_j \sum_k \beta_h^2(s_j) v_j^2 u_k^2 a_{kh'}^2. \end{aligned}$$

The second term of the last expression goes to zero as m and r go to infinity, because

$$\begin{aligned} \sum_j \beta_h^2(s_j) v_j^2 &\leq (\max v_j) \sum_j \beta_h^2(s_j) v_j, \\ \sum_k u_k^2 a_{kh'}^2 &\leq (\max u_k) \sum_k u_k a_{kh'}^2, \end{aligned}$$

and both sums converge to finite integrals. Therefore,

$$\lim_{\substack{m\to\infty \\ r\to\infty}} \mathrm{V}(y_i) = \iiiint \beta_h(s_1)\alpha_{h'}(t_1)\beta_h(s_2)\alpha_{h'}(t_2)\rho\{(s_1, t_1), (s_2, t_2)\}\mathrm{d}s_1\mathrm{d}s_2\mathrm{d}t_1\mathrm{d}t_2.$$

Then $\mathrm{V}(y_i)$ is bounded for any m and r, and a simple application of Tchebyshev's inequality implies that (4) goes to zero in probability as n goes to infinity, uniformly in m and r.

Regarding (5), note that

$$\beta_h(\mathbf{s})^\top \mathbf{V}\bar{\mathbf{Z}}\mathbf{U}\bar{\mathbf{Z}}^\top \mathbf{V}\beta_{h'}(\mathbf{s}) \leq \|\mathbf{U}^{1/2}\bar{\mathbf{Z}}^\top \mathbf{V}\beta_h(\mathbf{s})\| \|\mathbf{U}^{1/2}\bar{\mathbf{Z}}^\top \mathbf{V}\beta_{h'}(\mathbf{s})\|.$$

For a given index h, we can write $\mathbf{U}^{1/2}\bar{\mathbf{Z}}^\top \mathbf{V}\beta_h(\mathbf{s}) = \bar{\mathbf{w}}$, with $\mathbf{w}_i = \mathbf{U}^{1/2}\mathbf{Z}_i^\top \mathbf{V}\beta_h(\mathbf{s})$. The $\mathbf{w}_i$s are i.i.d. with $\mathrm{E}(\mathbf{w}_i) = 0$ and

$$\begin{aligned}
\mathrm{V}(w_{ik}) &= u_k \sum_j v_j \beta_h(s_j) \sum_{j'} v_{j'} \beta_h(s_{j'}) \mathrm{Cov}(z_{ijk}, z_{ij'k}) \\
&= u_k \sum_j v_j \beta_h(s_j) \sum_{j'} v_{j'} \beta_h(s_{j'}) \rho\{(s_j, t_k), (s_{j'}, t_k)\} \\
&\quad + \sigma^2 u_k \sum_j v_j^2 \beta_h^2(s_j).
\end{aligned}$$

As before, $\sum_j v_j^2 \beta_h^2 \to 0$ as $m \to \infty$, and since $\sum_k u_k = t_r - t_1$ is constant, we have

$$\lim_{\substack{m\to\infty \\ r\to\infty}} \sum_{k=1}^{r} \mathrm{V}(w_{ik}) = \iiint \beta_h(s_1)\beta_h(s_2)\rho\{(s_1, t), (s_2, t)\}\mathrm{d}s_1\mathrm{d}s_2\mathrm{d}t.$$

Since $\mathrm{E}(\|\bar{\mathbf{w}}\|^2) = n^{-1}\sum_{k=1}^{r} \mathrm{V}(w_{ik})$, a straightforward application of Markov's inequality implies that $\|\bar{\mathbf{w}}\|$ goes to zero in probability as n goes to infinity, uniformly in m and r, and consequently the same is true for (5). ■

2 Human mortality example: some individual plots

As explained in the paper, three countries are atypical: Finland, Spain and Italy. The corresponding mortality plots are shown in Figs. 1, 2 and 3.

3 Simulations

Figs. 4, 5 and 6 show boxplots of the simulated integrated squared errors $\|\hat{\mu} - \mu\|$ for the three models.

References

Gohberg, I., Goldberg, S. & Kaashoek, M. A. (2003). *Basic Classes of Linear Operators*. Basel: Birkhäuser Verlag.

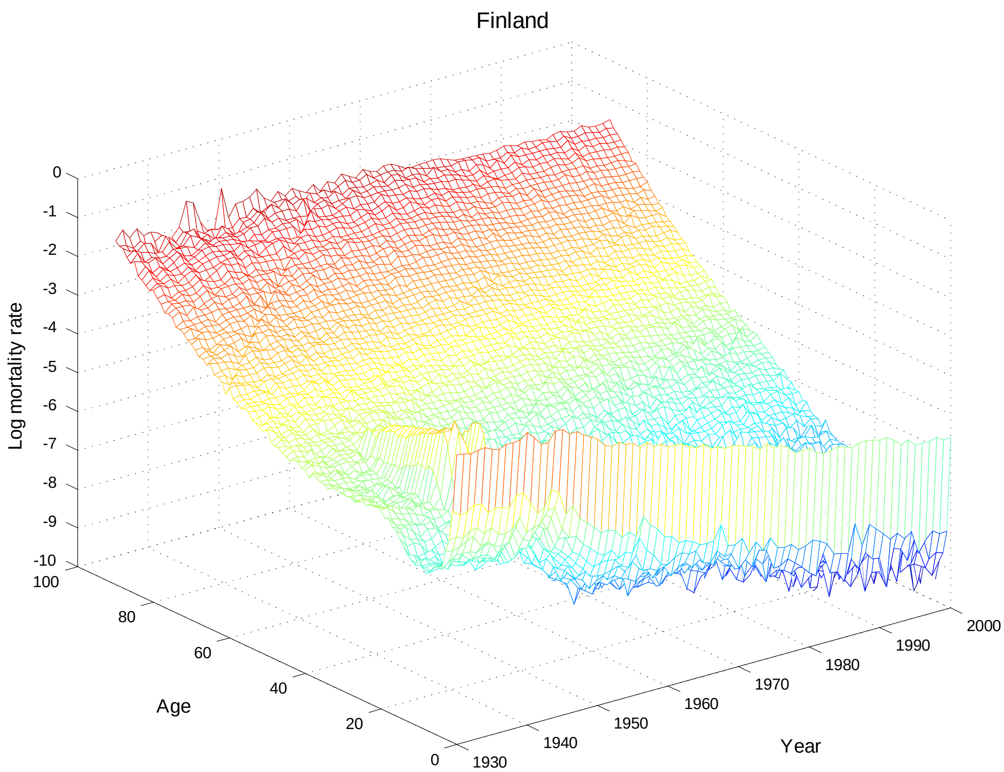

Figure 1: Human Mortality Data. Log-mortality rates for Finland.

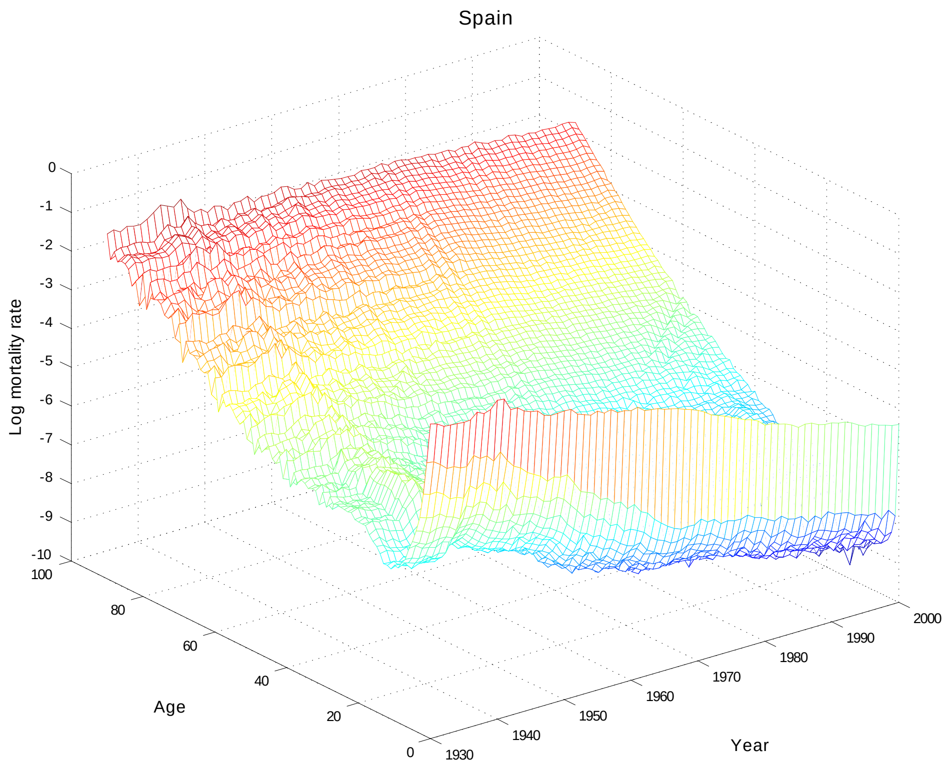

Figure 2: Human Mortality Data. Log-mortality rates for Spain.

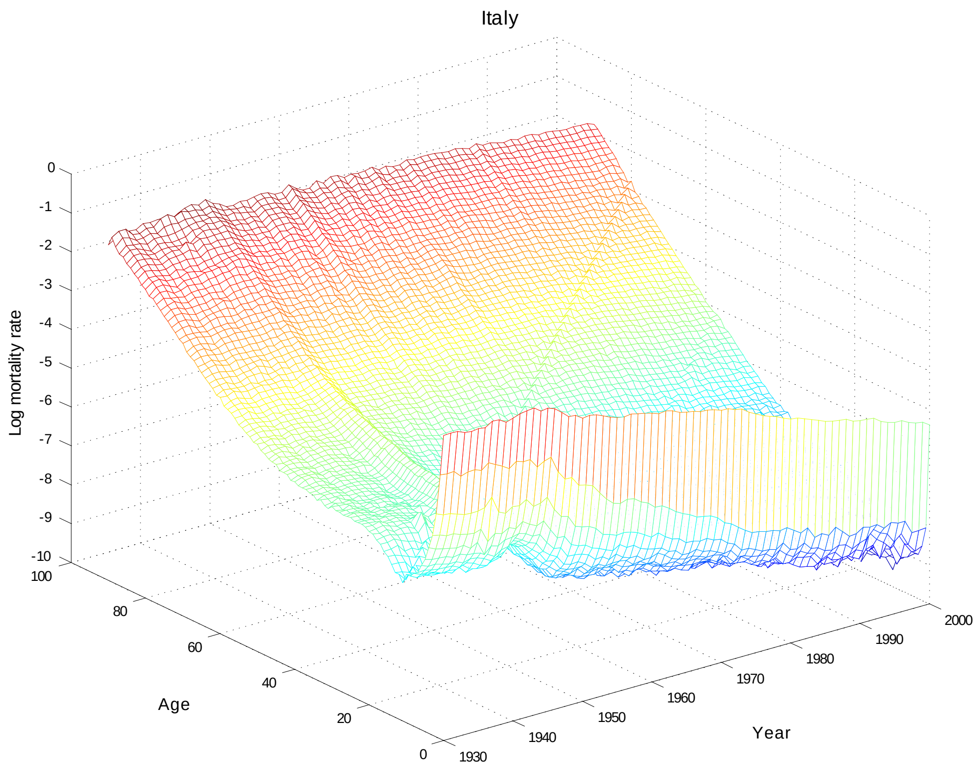

Figure 3: Human Mortality Data. Log-mortality rates for Italy.

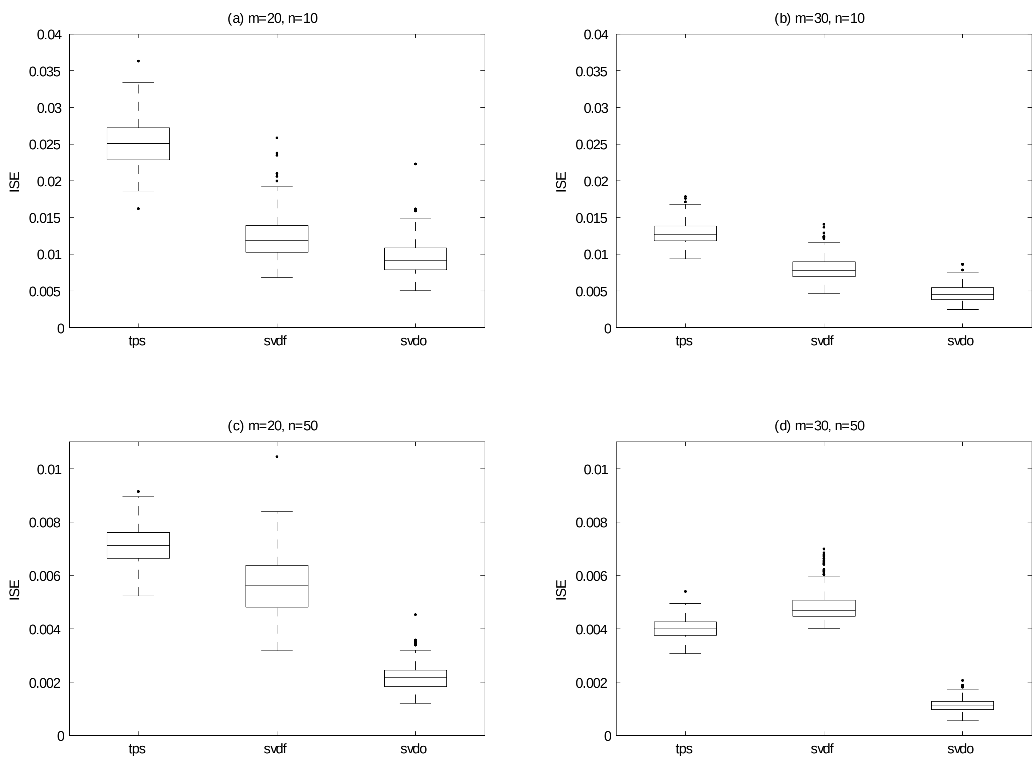

Figure 4: Simulation Results. Integrated squared errors for μ_1 and $\sigma = 1$.

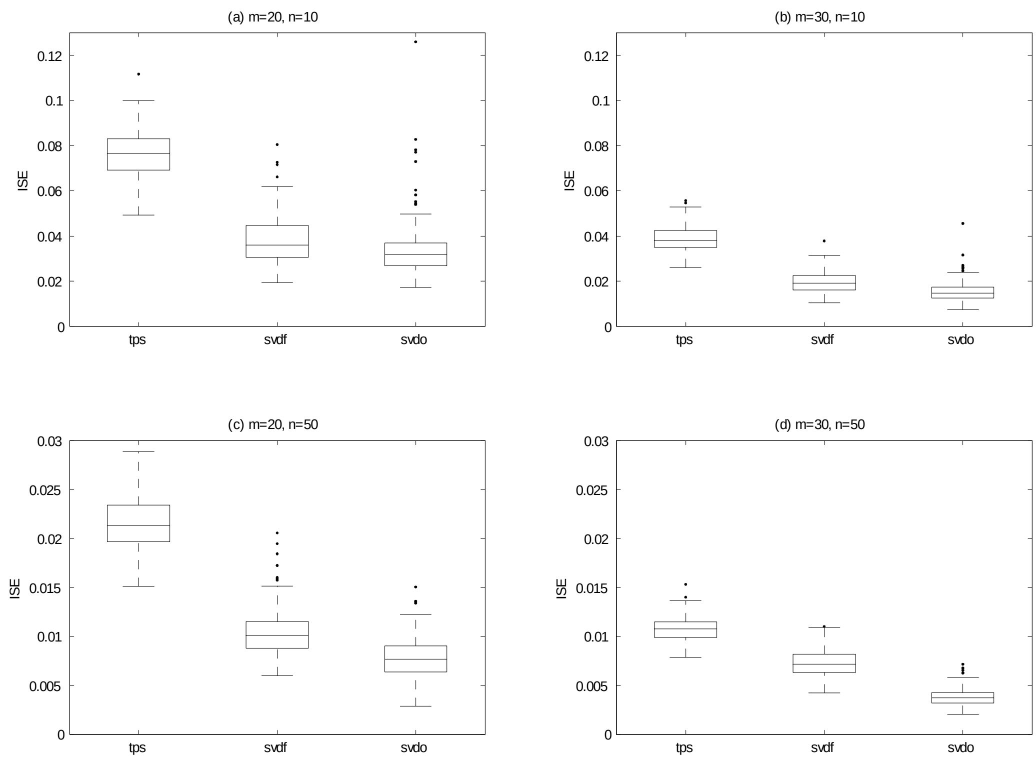

Figure 5: Simulation Results. Integrated squared errors for μ_1 and $\sigma = 2$.

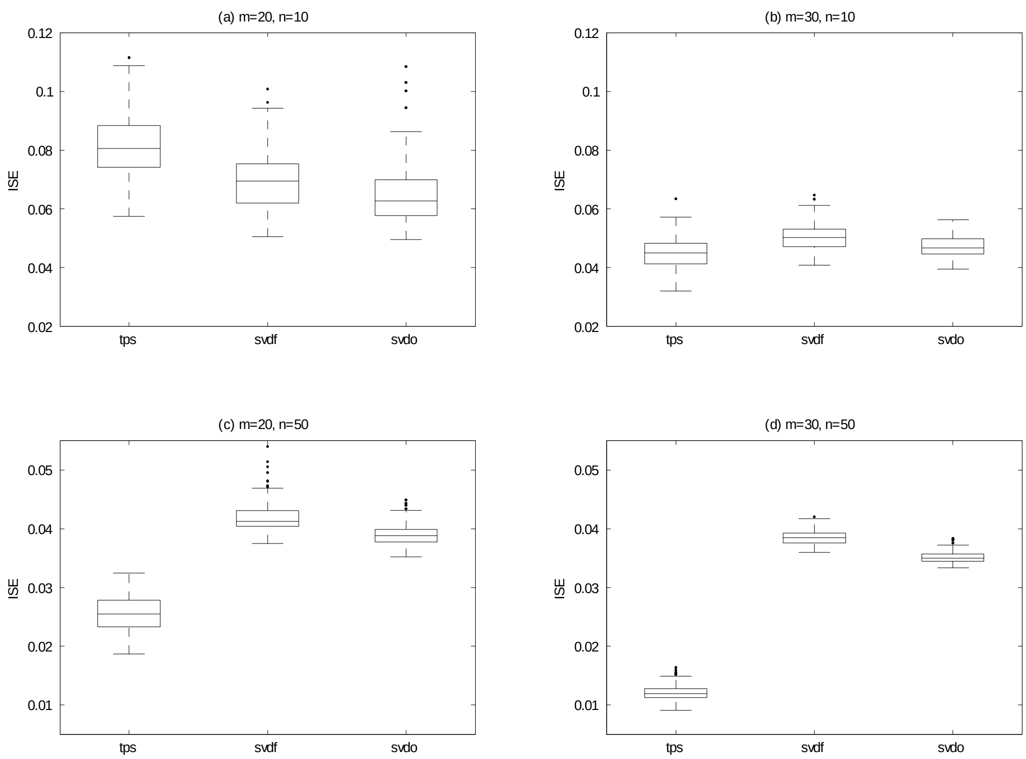

Figure 6: Simulation Results. Integrated squared errors for μ_2 and $\sigma = 2$.