



COLEGIO DE POSTGRADUADOS

INSTITUCIÓN DE ENSEÑANZA E INVESTIGACIÓN
EN CIENCIAS AGRÍCOLAS

CAMPUS MONTECILLO

POSTGRADO DE SOCIOECONOMÍA, ESTADÍSTICA E
INFORMÁTICA - ESTADÍSTICA

**Aplicaciones del Modelo LASSO Bayesiano en
Finanzas**

Yazmín García Salinas

T E S I S

PRESENTADA COMO REQUISITO PARCIAL PARA
OBTENER EL GRADO DE:

MAESTRA EN CIENCIAS


MONTECILLO, TEXCOCO, EDO. DE MÉXICO
2011

La presente tesis titulada: **Aplicaciones del Modelo LASSO Bayesiano en Finanzas**, realizada por la alumna: **Yazmín García Salinas**, bajo la dirección del Consejo Particular indicado ha sido aprobada por el mismo y aceptada como requisito parcial para obtener el grado de:

MAESTRA EN CIENCIAS

SOCIOECONOMÍA, ESTADÍSTICA E INFORMÁTICA ESTADÍSTICA

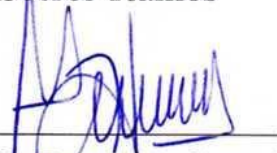
CONSEJO PARTICULAR

CONSEJERO 

Dr. José A. Villaseñor Alva

ASESOR 

Dr. David A. Sotres Ramos

ASESOR 

Dr. Gerardo H. Terrazas González

ASESOR 

Dr. Paulino Pérez Rodríguez

Aplicaciones del Modelo LASSO Bayesiano en Finanzas

Yazmín García Salinas

Colegio de Postgraduados, 2011

Resumen

En el presente trabajo se propone utilizar la metodología conocida como LASSO Bayesiano [Park y Casella (2008)] para ajustar un modelo de regresión lineal a indicadores económicos relacionados con el PIB (GDP: Gross Domestic Product por sus siglas en Inglés), utilizando como variable respuesta el PIB de Estados Unidos.

Esta técnica utiliza la ventaja de ver a LASSO como una estimación a posteriori cuando los parámetros de regresión tienen distribución a priori Laplace idénticamente distribuidas e independientes [Tibshirani (1996)], aprovechando la utilidad del muestreador de Gibbs [Casella (2001)] y los modelos jerárquicos [Lee (2004)]. Se utiliza el paquete BLR [de los Campos y Pérez (2010)] implementado en el programa *R* [R Development Core Team (2011)] para realizar una prueba de permutaciones para probar la significancia de los β 's. Se estima la correlación entre y y \hat{y} obtenido mediante el paquete BLR y otras metodologías como LASSO ordinario y regresión Ridge, utilizando a BLR como un modelo generalizado que incluye a LASSO Bayesiano presentado por Park y Casella (2008).

Palabras clave: Modelo lineal, Prueba de permutaciones, Muestreador de Gibbs, Modelos Jerárquicos.

Applications of The Bayesian LASSO Model in Finance

Yazmín García Salinas

Colegio de Postgraduados, 2011

Abstract

We propose to use Bayesian LASSO to fit a linear model to economic indicators related to the GDP (Gross Domestic Product), using as the response variable the U.S. GDP.

This technique uses the advantage of seeing LASSO as a posteriori estimate when the regression parameters are independent and identically distributed with Laplace distribution [Tibshirani (1996)], using the Gibbs sampling [Casella (2001)] and hierarchical models. BLR package [de los Campos y Pérez (2010)] implemented in the R [R Development Core Team (2011)], is used to perform a permutation test and verify the values of β . We estimate the correlation between y and \hat{y} obtained by the BLR package and others methodologies such as ordinary LASSO and Ridge regression, using the BLR as a generalized model including Bayesian LASSO described by Park y Casella (2008).

Key words: Linear model, Permutation test, Gibbs sampling, Hierarchical Models.

AGRADECIMIENTOS

Al Consejo Nacional de Ciencia y Tecnología (CONACYT) por el apoyo económico brindado durante la realización de mi maestría.

Al Colegio de Postgraduados, en particular al Programa de Estadística, por haberme brindado la oportunidad de continuar mi formación académica.

A los integrantes de mi Consejo Particular:

Al Dr. José Aurelio Villaseñor Alva mi mayor agradecimiento, por su interminable paciencia, por sus consejos y ayuda en la realización del presente trabajo de tesis.

Al Dr. Gerardo H. Terrazas González, por su orientación, apoyo y colaboración desinteresada en el presente trabajo.

Al Dr. Paulino Pérez Rodríguez, por sus acertadas sugerencias, ayuda incondicional y por dedicar parte de su tiempo en asesorías para terminar este trabajo de tesis.

Al Dr. David A. Sotres Ramos, por su colaboración desinteresada en el presente trabajo.

A mis profesores, compañeros de clase, personal administrativo y a todas aquellas personas que de alguna manera fueron coparticipes de esta tarea, a todos gracias.

DEDICATORIA

A mis Padres Venancio García Cruz y Cirila Salinas Hernandez.

A mis Hermanas Ana Julieta y Bertha Cristina.

A mis Hermanos Jose Venancio y Carlos.

A Levi Cortés Cortés.

Gracias por todo su apoyo...

Índice

1. Introducción	1
2. Objetivos	4
2.1. Objetivos Generales	4
2.2. Objetivos particulares	4
3. Marco Teórico	5
3.1. Regresión Ridge	7
3.2. LASSO	10
4. LASSO Bayesiano	12
4.1. Modelo Jerárquico	12
4.2. Elección del parámetro de regularización en el LASSO Bayesiano	13
4.2.1. Estimación de λ mediante la maximización de la verosimilitud marginal	14
4.2.2. Hiper-aprioris para el parámetro de LASSO	14
4.3. Algoritmo de Gibbs	15
4.4. Regresión Lineal Bayesiana	16
4.5. Cadenas de Markov Monte Carlo con Saltos Reversibles	17

Índice

5. Prueba de Permutaciones	18
5.1. Remuestreo	18
5.2. Prueba de permutaciones	18
6. Aplicación a Indicadores Económicos	21
6.1. Definición de variables	21
6.2. Ajuste del modelo	21
6.3. Aplicando la prueba de permutaciones	23
6.4. Blasso mediante saltos reversibles	26
7. Discusión de Resultados	31
8. Conclusiones y Recomendaciones	34
Referencias	34
Anexos	37
Anexo A: Distribución Doble Exponencial	37
Anexo B: Actualización en Bayes Epírico	39
Anexo D: Rutinas en R	40

Índice de tablas

6.1. Variable respuesta y covariables.	22
6.2. Variable respuesta y covariables (Continuación).	23
6.3. Valores de $ \hat{\beta}_j > \mathbf{Q}_{95}$ obtenidos para el modelo LASSO Bayesiano	25
6.4. Variables con parámetro $ \hat{\beta}_j > \mathbf{Q}_{95}$	25
6.5. Probabilidad de $\beta_j \neq 0$	27
6.6. Variables con parámetro $\beta_j \neq 0$	27
6.7. Información de los parámetros con $P(\beta_j \neq 0) > 0.7$ obtenidos en R utilizando el paquete Blasso	28
6.8. Correlación entre y y \hat{y}	28
7.1. Variables finales obtenidas en el ajuste del modelo	31
7.2. Media (BLASSO) y Coeficiente del parámetro (BLR) obtenidas en el ajuste del modelo	33

Índice de figuras

4.1. Densidad de una distribución normal y doble exponencial (ambas con media cero y varianza igual a uno).	16
6.1. Efecto estimado de las covariables en el modelo $y = X\beta + \epsilon$ obtenidos utilizando LASSO Bayesiano.	24
6.2. Media de $\hat{\beta}_j$, cada uno con cuantil del 2.5% al 97.5% representado por una línea punteada.	24
6.3. Distribución empírica de la estadística de prueba.	25
6.4. Coeficientes de Lasso Bayesiano utilizando el paquete Blasso.	26
6.5. Diagrama de cajas para β_j en el modelo Blasso.	27
6.6. Probabilidad de $\beta_j \neq 0$	28
6.7. Grafica de y vs \hat{y} utilizando la metodología LASSO Bayesiano.	29
6.8. Grafica de y vs \hat{y} utilizando la metodología LASSO Clásico.	29
6.9. Grafica de y vs \hat{y} utilizando la regresión Ridge.	30
6.10. Grafica de y vs \hat{y} utilizando el paquete Blasso.	30

Capítulo 1

Introducción

El significado de finanzas nos indica que la actividad que la comprende es la ciencia correspondiente a la administración de dinero, ya que todos los individuos así como las entidades empresariales y las organizaciones suelen obtener dinero para que el mismo sea invertido o gastado, de aquí se tiene que el conocimiento de las finanzas nacionales e internacionales ayuda de dos maneras muy importantes. Primero, a decidir la manera en que los eventos internacionales afectarán a una empresa u organización y cuáles son los pasos que pueden tomarse para explotar los desarrollos positivos para aislar a la organización de los daños que pudieran existir.

Se define finanzas de la siguiente manera: Las finanzas, en economía, son las actividades relacionadas con los flujos de capital y dinero entre individuos, empresas o Estados. Por extensión también se denomina finanzas el estudio de esas actividades como especialidad de la Economía y la administración que estudia la obtención y gestión, por parte de una compañía, individuo o del Estado, de los fondos que necesita para cumplir sus objetivos y de los criterios con que dispone de sus activos. En otras palabras, estudia lo relativo a la obtención y gestión del dinero y de otros valores como títulos, bonos, etc. Las finanzas tratan, por lo tanto, de las condiciones y oportunidad en que se consigue el capital, de los usos de éste y de los pagos e intereses que se cargan a las transacciones en dinero, es decir, las finanzas son el estudio de los flujos de efectivo.

El concepto ampliado de finanzas es el de una ciencia que, utilizando modelos matemáticos, brinda las herramientas para optimizar los recursos materiales de las empresas, las personas o el Estado. La definición de finanzas comprende varios factores que se encuentran involucrados directamente con ellas, por ejemplo, las inversiones, los corredores de bolsa, la planificación financiera personal, los planificadores y asesores financieros, los gerentes de carrea, los analistas de valores, los agentes de bienes raíces, etc. A esta altura se debe comprender que las finanzas son la base de toda actividad comercial, y sin ellas la misma no se podría llevar a cabo, por ello es bueno tener en cuenta que el concepto de finanzas, no solo se limita a explicar de qué se tratan, sino que también señalar

1. Introducción

todos los factores que en ellas inciden como los indicadores de las finanzas nacionales e internacionales.

Entre los eventos que afectan a las empresas o Estados, se encuentran la variación de los tipos de cambio, en las tasas de interés, como en las tasas de inflación y en los valores de los activos. Debido a los estrechos vínculos que existen entre los mercados, los eventos en territorios distintos tienen efectos que se dejan sentir inmediatamente en todo el planeta. Esta situación ha hecho que para los administradores, tanto los formados como los aspirantes, sea un imperativo, hacer una detallada revisión del excitante y dinámico campo de las finanzas internacionales y nacionales [Baumohl (2008)].

Las finanzas nacionales por su parte están altamente relacionadas con algunos indicadores económicos que pueden dar información de gran importancia de hacia dónde va la economía de un país. Un indicador económico es una relación de las cifras extractadas de los estados financieros y demás informes del país u organización con el propósito de formarse una idea acerca de su comportamiento; se entienden a un indicador como la expresión cuantitativa del comportamiento o el desempeño de toda una organización o una de sus partes, cuya magnitud al ser comparada con algún nivel de referencia, puede estar señalando en donde se debe invertir como Estado, o en dónde se encuentra una desviación sobre la cual se tomaran acciones correctivas o preventivas según el caso. La interpretación de los resultados que arrojan los indicadores económicos y financieros está en función directa a las actividades y controles internos de las empresas u organizaciones como también a los períodos cambiantes causados por los diversos agentes internos y externos que las afectan.

La importancia de conocer la relación entre los indicadores económicos es básica, pues cada uno de ellos es un barómetro que mide la actividad en un segmento particular de la economía, dando pie a conocer hacia donde se dirigen los posibles cambios y repercusiones que puede generar en el país. Gobiernos y grupos privados de algunos países realizan más de una docena de reportes semanales, mensuales o trimestrales que presentan básicamente en esos periodos [Baumohl (2008)]. Estos reportes de indicadores sobre ventas, compras, créditos solicitados, hipotecas etc. están relacionados directamente con la dirección de la economía mundial.

Uno de los actuales retos de la generación de estadísticas de un país es contar con información relevante y oportuna para la toma de decisiones y seguimiento de los compromisos internacionales del mismo. Los indicadores económicos, sin duda, constituyen una de las herramientas indispensables para el logro de estas acciones, el notar sus relaciones será entonces de gran apoyo en la toma de mejores decisiones económicas para una organización. En los estudios financieros es necesario analizar la relación entre los indicadores económicos y para algunos de ellos sólo se cuenta con información de los periodos inmediatos, es decir, no se tiene la información necesaria para llevar a cabo una regresión simple [Gramacy y Pantaleo (2010)], si el caso lo requiere, esta es la principal razón de la búsqueda de alternativas a los procesos clásicos de regresión, pues infinidad de problemas de este tipo se presentan a diario en finanzas [Gramacy *et al.* (2009)], por ejemplo,

1. Introducción

el caso donde algún indicador económico depende de un conjunto de otros indicadores económicos, de los cuales no se tiene suficiente información de los periodos que se requieren, aquí es entonces, donde se presenta el problema principal, es decir, el número de variables de las cuales depende este indicador es mayor que el número de observaciones disponibles, además predictores redundantes afectan a la precisión de las estimaciones de los coeficientes. Este último aspecto es verdaderamente importante en el caso de variables correlacionadas, una realidad a tener presente en los datos económicos.

Ya se han propuesto métodos alternativos a la regresión para la solución de este problema, algunos de estos métodos se revisan en el capítulo 3. Otra alternativa se basa en la aplicación de técnicas bayesianas [Lee (2004)]. En el presente trabajo se desarrolla una aplicación del Método LASSO Bayesiano en datos Financieros, utilizando la metodología propuesta por Park y Casella (2008), la cual se basa en estimaciones bayesianas utilizando métodos iterativos como el muestreo de Gibbs [Casella (2001)] dentro de los modelos jerárquicos. También se utilizan datos reales de indicadores económicos propuestos por Baumohl (2008) para realizar el ajuste. La estimación propuesta es comparada con la de otros procedimientos que se utilizan con la finalidad de obtener la mejor estimación a partir de su comparación. Algunas aplicaciones de LASSO y LASSO Bayesiano en finanzas se pueden encontrar en Gramacy *et al.* (2009) y Gramacy y Pantaleo (2010).

Capítulo 2

Objetivos

2.1. Objetivos Generales

- Aplicar la metodología LASSO Bayesiano a datos financieros en donde el número de variables explicativas es mayor al número de observaciones (variables independientes).

2.2. Objetivos particulares

- Utilizar los indicadores económicos apoyándose en la propuesta de Baumohl (2008) para llevar a cabo un análisis de regresión lineal entre el PIB y sus principales indicadores.
- Aplicar el Método Lasso Bayesiano.
- Aplicar la metodología para hallar el modelo que relaciona un indicador económico con un conjunto de variables (otros indicadores económicos de los cuales depende).
- Utilizar estos resultados para predecir y observar la dirección que puede tomar la economía del país referencia.

Capítulo 3

Marco Teórico

Los indicadores económicos son valores estadísticos obtenidos a través del análisis comparativo entre un año y otro de un periodo determinado que reflejan el comportamiento de las principales variables económicas, financieras y monetarias. La relación que existe entre un indicador y otras variables puede explicarse mediante un modelo de regresión lineal múltiple.

En estadística el modelo de regresión clásico esta dado por:

$$y = X\beta + \epsilon \quad (3.1)$$

donde y es el vector de respuestas, en este caso, el indicador que depende de otras variables u otros indicadores económicos, X es la matriz de covariables, β son los parámetros que deben hallarse para completar el modelo o relación lineal entre la variable independiente y y el conjunto de covariables X , por último, ϵ es el vector de errores con distribución normal multivarida con media 0 y varianza desconocida $\sigma^2 I$, donde I es la matriz idéntica.

El problema de la regresión consiste en elegir valores determinados para los parámetros desconocidos β_k con $k = 1, \dots, p$, de modo que la ecuación quede completamente especificada. Para ello se necesita un conjunto de observaciones. En una observación cualquiera i -ésima ($i = 1, \dots, n$) se registra el comportamiento simultáneo de la variable dependiente y las variables explicativas (las perturbaciones aleatorias se suponen no observables).

Utilizando mínimos cuadrados, el estimador del vector de parámetros β es :

$$\hat{\beta} = (X'X)^{-1}X'y \quad (3.2)$$

3. Marco Teórico

obtenido al minimizar

$$\phi(\beta) = (y - X\beta)'(y - X\beta), \quad (3.3)$$

dada por β que minimiza la suma del cuadrado de los residuales.

Es preciso saber que para llevar a cabo un análisis de regresión lineal por el método de mínimos cuadrados, es necesario que se cumpla con los siguientes supuestos:

- La relación entre las variables es lineal.
- Los errores en la medición de las variables explicativas son independientes entre sí.
- Los errores tienen varianzas constantes.
- Los errores tienen media cero (los errores de una misma magnitud y distinto signo son equiprobables).
- El error total es la suma de todos los errores.
- $n > p + 1$ donde n es el número de observaciones y p es el número de covariables, en caso contrario no se dispone de información suficiente para estimar los parámetros del modelo.
- Las variables regresoras x_1, x_2, \dots, x_p son linealmente independientes.

En el caso de los indicadores económicos es muy poco probable que se cumplan los supuestos del modelo de regresión dado que se tienen más covariables que observaciones, puesto que muchas veces no se cuenta con los periodos necesarios para llevar a cabo una regresión lineal con métodos clásicos como Mínimos Cuadrados. Existen métodos alternativos que pueden ser de gran ayuda cuando las covariables no son linealmente independientes, es decir, existe correlación entre variables dado que $p > n$ y existe la condición de multicolinealidad.

La colinealidad es un problema del análisis de regresión que consiste en que los predictores del modelo están relacionados constituyendo una combinación lineal. Este hecho tiene consecuencias fundamentales en el modelo de regresión: si los predictores se encuentran en combinación lineal, la influencia de cada uno de ellos en el criterio no puede distinguirse al quedar solapados unos con otros; no se consigue una explicación del fenómeno en cuestión; los pronósticos no son nada fiables, puesto que otra combinación de predictores introducida en el modelo variando el orden, produce predicciones en el criterio contradictorias; no se realiza una selección adecuada del orden de entrada de los regresores en el modelo, y un largo etcétera. Es un problema que no tiene fácil solución, ya que en

3.1. Regresión Ridge

definitiva se trata de pedirle a la muestra de datos más información de la que posee. Para corregirla sólo cabe actuar en alguno de los siguientes sentidos:

Eliminar variables predictoras, con lo que se reduce el número de parámetros a estimar o incluir información externa a los datos originales. Si se opta por el primero de ellos, se trata de suprimir, o bien ciertas variables que se encuentren altamente correlacionadas, o bien algunas combinaciones lineales mediante el análisis de componentes principales aplicado a la regresión. La segunda alternativa conduce a trabajar con estimadores contraídos o bayesianos. En ambas opciones se sustituyen los estimadores mínimo cuadráticos de los coeficientes de regresión por estimadores sesgados pero con menor varianza que los de mínimos cuadrados. Estos procedimientos forman parte de la regresión sesgada, no lineal, pero que sigue cumpliendo el supuesto de los mínimos cuadrados. Además, estos estimadores, a pesar de ser sesgados, tienen un error cuadrático medio mucho menor, que es lo que se pretende para corregir la colinealidad.

Básicamente hasta hace algunos años sólo había tres propuestas importantes: Regresión Ridge [Hoerl y Kennard (1970)], Componentes principales [Hotelling (1933)] y Mínimos Cuadrados Parciales [Wold (1975)]. Sin embargo el problema de multicolinealidad también está relacionado con los métodos de selección de variables y esto puede ser considerado como una cuarta manera de resolver el problema de multicolinealidad. Por último puede ser considerado el análisis bayesiano como una nueva alternativa para resolver el problema cuando $p \gg n$.

En este trabajo se revisaran algunos métodos alternativos de regresión muy utilizados como la regresión Ridge y algunos más recientes como la estimación LASSO y LASSO Bayesiano, siendo este último el método propuesto para obtener la estimación de los parámetros de regresión en el modelo.

3.1. Regresión Ridge

El estimador de Mínimos Cuadrados (MC), es un caso especial del estimador general Ridge propuesto por Hoerl y Kennard (1970), para ajustar un modelo de regresión polinomial.

La estimación de un modelo de regresión lineal por mínimos cuadrados se basa en buscar el estimador de mínima varianza en la clase de los estimadores insesgados, de esta forma se obtiene el estimador de menor error cuadrático medio dentro de los de esta clase. Un método alternativo de estimación consiste en buscar estimadores que pueden ser sesgados pero con menor error cuadrático medio (ECM). Los métodos de regresión contraída parten de esta idea y tratan de minimizar el $ECM(\hat{\beta})$. Se debe tener en cuenta que al comparar los estimadores de mínimos cuadrados con los estimadores de regresión contraída no siempre se están comparando estimadores insesgados con estimadores sesgados ya que los

3.1. Regresión Ridge

estimadores por mínimos cuadrados son insesgados si el modelo de regresión en estudio es el adecuado en donde se cumplan las hipótesis básicas.

La estimación por mínimos cuadrados ordinarios es una técnica de análisis numérico que intenta minimizar la suma de cuadrados de las diferencias ordenadas (llamadas residuos o errores) entre los puntos generados por la función y los correspondientes en los datos. Dados los supuestos del modelo clásico de regresión lineal, los valores estimados de mínimos cuadrados poseen algunas propiedades ideales u óptimas. Estas propiedades están contenidas en el teorema conocido como teorema Gauss-Markov y son las siguientes:

- Los estimadores de mínimos cuadrados son combinaciones lineales.
- Insesgados
- De varianza mínima.

La Regresión Ridge supone un procedimiento de ajuste de los estimadores mínimo cuadráticos con el objetivo de reducir su variación. Constituye toda una alternativa a la estimación mínimo cuadrática, y además proporciona una evidencia gráfica de los efectos de la colinealidad en la estimación de los coeficientes de regresión. Dicho procedimiento se encuentra dentro del grupo de las regresiones parciales o sesgadas, consideradas como no lineales. Una de las consecuencias de la colinealidad es el aumento de la varianza asociada a la estimación de los coeficientes de regresión, con lo que los coeficientes estimados pueden no brindar la suficiente confianza o incluso ser erróneos si los vectores predictores no son ortogonales. La Regresión Ridge proporciona unos nuevos estimadores que garantizan precisamente minimizar dicha varianza, de forma que constituye un método alternativo cuando las variables explicativas son altamente no ortogonales.

El nombre de Regresión Ridge se debe a los trabajos desarrollados por Hoerl (1964) al aplicar el Análisis Ridge a las superficies de respuesta de segundo orden que resultan de utilizar muchos regresores. Dicho análisis Ridge y la regresión Ridge presentan procesos matemáticos similares. La primera exposición detallada de esta regresión se encuentra en el artículo de Hoerl y Kennard (1970), donde se contempla el proceso de cálculo de los estimadores Ridge, su expresión canónica, el gráfico de trazo Ridge y el análisis gráfico. Posteriormente, Marquardt y Snee (1975) llevan a cabo una buena aplicación de todos estos procesos con un estudio detallado sobre la interpretación y validación del modelo. Tal y como se encuentra ya desarrollada esta técnica, supone un aproximación que responde a dos problemas: (a) detección, y (b) estimación de la multicolinealidad, cuando hay sospechas de su existencia.

En general, la filosofía que subyace en el análisis es la comparación entre la estimación mínimo cuadrática común, MCO (Mínimos Cuadrados Ordinarios), y la estimación mediante la estrategia del trazo Ridge. Ambas alternativas son enfrentadas y son estudiadas en relación a lo que sería la verdadera estimación de los coeficientes (en términos de esperanzas) que es desconocida. En este proceso de comparaciones y estudio de propiedades

3.1. Regresión Ridge

se llega a demostrar que en los casos de colinealidad, los estimadores resultantes por el trazo ridge se encuentran más próximos o concentrados a los verdaderos valores de los parámetros de regresión, y por lo tanto dichas estimaciones presentan una varianza menor a la que poseen las estimaciones por MCO (incrementada por el efecto de la colinealidad). Estos nuevos estimadores ridge son parciales o sesgados, pero dado que su media cuadrática es menor, son más estables, en el sentido de que no van a estar afectados por las variaciones de los datos (la demostración de que los estimadores ridge se distribuyen con una media cuadrática de error menor se encuentra en Hoerl y Kennard (1970)). El criterio para decidir cuándo los estimadores ridge son preferibles a los estimadores MCO depende de los valores de los verdaderos coeficientes del modelo de regresión, pero como éstos son desconocidos, la regresión ridge sería útil cuando se sospecha de una extrema colinealidad.

Los estimadores ridge de los coeficientes de regresión se pueden hallar alterando ligeramente las ecuaciones normales. Esta modificación consiste en añadir pequeñas cantidades positivas a la diagonal de la matriz de datos siempre y cuando ésta se encuentre en forma estandarizada. Tomando entonces la forma estandarizada del modelo en (3.1)

Las ecuaciones de estimación para el modelo ridge con p variables explicativas son:

$$\begin{aligned}(1 + \lambda)\beta_1 + r_{12}\beta_2 + \dots + r_{1p}\beta_p &= r_{1y} \\ r_{12}\beta_1 + (1 + \lambda)\beta_2 + \dots + r_{2p}\beta_p &= r_{2y} \\ &\vdots \\ r_{1p}\beta_1 + r_{2p}\beta_2 + \dots + (1 + \lambda)\beta_p &= r_{py}\end{aligned}$$

donde r_{ij} es la correlación simple entre los predictores i y j , y r_{iy} es la correlación entre la i -ésima variable explicativa y la variable respuesta. La solución a las ecuaciones anteriores es el grupo de coeficientes de regresión ridge estimados:

$$\tilde{\beta}_1^*, \tilde{\beta}_2^*, \dots, \tilde{\beta}_p^*$$

es decir, la estimación esta basada en la matriz $[X'X + \lambda I_p]$. A. E. Hoerl sugirió en 1962 que para manejar la inflación de varianza y la inestabilidad general asociada a la estimación de mínimos cuadrados, se puede utilizar:

$$\tilde{\beta}^* = [X'X + \lambda I]^{-1} X'y \tag{3.4}$$

para $\lambda \geq 0$ ó

$$\tilde{\beta}^* = WX'y \tag{3.5}$$

3.2. LASSO

Como puede verse, el parámetro fundamental que distingue la regresión ridge de la regresión MCO es λ , denominado estimador ridge. Nótese que cuando λ es cero, los β_i coinciden con los estimadores de MCO (sería el caso de ortogonalidad de la matriz de datos X). Mientras que en MCO los estimadores de los parámetros se definen como en la ecuación 3.2.

3.2. LASSO

LASSO(Least Absolute Shrinkage and Selection Operator de sus siglas en Inglés) es una versión restringida de la estimación por mínimos cuadrados. Esta se obtiene de una simple modificación al algoritmo LAR [Efron *et al.* (2004)] descrito como un procedimiento por etapas, comenzando en cero coeficientes, LASSO actualiza las estimaciones hasta terminar con el ajuste de mínimos cuadrados de todos los coeficientes, siendo lineal por segmentos en las diferentes etapas.

Sean los predictores o regresores x_1, x_2, \dots, x_p vectores de dimensiones $n \times 1$ y sea y un vector de dimensión $n \times 1$, la variable dependiente o respuesta. Sin pérdida de generalidad se puede suponer que los predictores han sido normalizados, para tener media cero y norma uno, es decir:

$$\sum y_i = 0; \sum x_{ij} = 0; \sum x_{ij}^2 = 1; j = 1, \dots, p$$

. Sea el conjunto de estimadores

$$\hat{\beta} = (\hat{\beta}_1, \hat{\beta}_2, \dots, \hat{\beta}_p)'$$

Además sea \hat{y} la predicción del vector de respuestas, dado por:

$$\hat{y} = \sum x_j \hat{\beta}_j = X \hat{\beta}$$

Siendo X la matriz de regresores de dimensión $n \times p$. Llamando \hat{u} al vector de residuos : $\hat{u} = y - \hat{y}$ es posible escribir la suma de residuos al cuadrado como:

$$S(\beta) = \|y - \hat{y}\|^2 = \sum (y_i - \hat{y}_i)^2 = \hat{u}'\hat{u}$$

Denominando $T(\hat{\beta}) = \sum |\hat{\beta}_i|$ a la norma absoluta de la suma de los coeficientes estimados, entonces, LASSO selecciona los $\hat{\beta}$ minimizando la suma de los residuos al cuadrado, sujeta a la condición de que

$$T(\hat{\beta}) \leq t$$

3.2. LASSO

siendo t un valor conocido. Este problema puede ser visto como:

$$\min_{\beta} \left\{ \sum (y_i - x_i' \beta)^2 \text{ sujeto a } \sum_j |\beta_j| \leq t \right\}, \quad (3.6)$$

lo cual usando multiplicadores de Lagrange es equivalente a:

$$\min_{\beta} \left\{ (\tilde{y} - X\beta)^T (\tilde{y} - X\beta) + \lambda \sum_{j=1}^p |\beta_j| \right\} \text{ donde } \tilde{y} = y - \bar{y} \mathbf{1}_n \quad (3.7)$$

La solución a este problema recurre a la programación cuadrática [Efron *et al.* (2004)].

El parámetro λ funciona como parámetro de regularización, jugando un papel central: si este es cero, la solución en 3.7 es la solución de mínimos cuadrados ordinarios.

Por construcción la versión no bayesiana de LASSO admite a lo más $n - 1$ coeficientes de regresión diferentes de cero [Park y Casella (2008)], esto algunas veces no es deseable en los modelos de regresión, pues no existe razón para que el número de regresores distintos de cero tenga que ser más pequeño que el número de observaciones.

Capítulo 4

LASSO Bayesiano

4.1. Modelo Jerárquico

Tibshirani (1996) propuso utilizar la suma de los valores absolutos de los parámetros de regresión como la penalización en la estimación de mínimos cuadrados, que simultáneamente produce selección de variables y reducción de coeficientes, a esta metodología se le llamo LASSO. Es bien sabido que la solución en (3.7) puede implicar la reducción a cero de algunos elementos de β [Tibshirani (1996)]. Una de las formas de notar esto es al considerar la solución Bayesiana de LASSO. Se sugiere que la estimación LASSO puede ser interpretada como un modo de estimación a posteriori cuando los parámetros de regresión tienen distribuciones a prioris Laplace (i.e., doble exponencial) idénticamente distribuidas e independientes. En Park y Casella (2008) se considera este análisis utilizando una distribución a priori condicional Laplace especificada de la forma:

$$\pi(\beta \mid \sigma^2) = \prod_{j=1}^p \frac{\lambda}{2\sqrt{\sigma^2}} \exp^{-\lambda|\beta_j|/\sqrt{\sigma^2}} \quad (4.1)$$

y la apriori marginal no informativa de escala invariante

$$\pi(\sigma^2) \propto \frac{1}{\sigma^2}$$

Condicionar en σ^2 es importante ya que asegura que las distribuciones posteriores sean unimodales, ya que la falta de unimodalidad retrasa la convergencia del muestreo de Gibbs y hace que los puntos de estimación sean menos significativos.

Usando el hecho de que la densidad Laplace se puede escribir como una mezcla de normales, podemos escribir el problema original como un modelo jerárquico, cuya estructura facilita la implementación del muestreador de Gibbs.

4.2. Elección del parámetro de regularización en el LASSO Bayesiano

$$\frac{a}{2} \exp^{-a|z|} = \int_0^\infty \frac{1}{\sqrt{2\pi s}} \exp^{-z^2/2s} \frac{a^2}{2} \exp^{-a^2 s/2} ds, \quad a > 0 \quad (4.2)$$

Lo que sugiere la representación jerárquica de modelo completo de la siguiente manera:

$$\begin{aligned} y | X, \beta, \sigma^2 &\sim N_n(X\beta, \sigma^2 I), \\ \beta | \sigma^2, \tau_1^2, \dots, \tau_p^2 &\sim N_p(0_p, \sigma^2 D_\tau), \\ D_\tau &= \text{diag}(\tau_1^2, \dots, \tau_p^2), \\ \sigma^2, \tau_1^2, \dots, \tau_p^2 &\sim \pi(\sigma^2) d\sigma^2 \prod_{j=1}^p \frac{\lambda^2}{2} \exp^{-\lambda^2 \tau_j^2/2} d\tau_j^2, \\ \sigma^2, \tau_1^2, \dots, \tau_p^2 &> 0 \end{aligned} \quad (4.3)$$

En Park y Casella (2008) se utiliza la a priori impropia $\pi(\sigma^2) \propto 1/\sigma^2$, aunque con cualquier gamma invertida se obtienen distribuciones conjugadas. Las distribuciones para β, σ^2 y $\tau_1^2, \dots, \tau_p^2$ siguen siendo fáciles de muestrear, sólo dependiendo del vector de respuestas centrado \tilde{y} .

El papel que juegan los parametros $\tau_1^2, \dots, \tau_p^2$ es visiblemente más claro mediante un cambio de variable en 4.3, utilizando $\beta_j^{(*)} = \tau_j^{-1} \beta_j$ en lugar de β_j , después de este cambio de variable el producto de la función de verosimilitud y la distribución a priori condicional de los coeficientes de regresión $N_n(y|X\beta, \sigma^2 I) N_p(\beta|\sigma^2 D_\tau)$ se convierten en $N_n(y|X^{(*)}\beta^{(*)}, \sigma^2 I) N_p(\beta^{(*)}|\sigma^2 I)$, donde $X^{(*)} = X D_\tau$ y $\beta^{(*)}$ es un vector de coeficientes de regresión con varianza homogénea. Así, una forma de ver esta clase de modelos de regresión es como un modelo estandar de regresión bayesiana con parámetros adicionales desconocidos D_τ , los cuales asignan diferentes pesos a las columnas de X con $\tau_j \rightarrow 0$ siendo equivalente a la eliminación de la j -ésima covariable del modelo [de los Campos *et al.* (2009)].

4.2. Elección del parámetro de regularización en el LASSO Bayesiano

Es posible obtener el parámetro de Lasso ordinario mediante validación cruzada, validación cruzada generalizada o en ideas de estimación del riesgo insesgado de Stein [Tibshirani (1996)]. En la estimación LASSO Bayesiano es posible obtener el parámetro mediante Bayes empírico a través de la maximización de la verosimilitud marginal [Casella (2001)] o utilizando una a priori adecuada.

4.2. Elección del parámetro de regularización en el LASSO Bayesiano

4.2.1. Estimación de λ mediante la maximización de la verosimilitud marginal

En Casella (2001) se propone un algoritmo EM Monte Carlo que complementa al muestreo de Gibbs y proporciona estimaciones de los hiperparámetros a través de la maximización de la verosimilitud marginal. Para LASSO Bayesiano cada iteración del algoritmo involucra una corrida del muestreo de Gibbs utilizando el valor estimado de λ de la iteración previa. Específicamente la iteración k utiliza el muestreo de Gibbs y el modelo en (4.3) con hiperparámetro $\lambda^{(k-1)}$,

$$\lambda^{(k)} = \sqrt{\frac{2p}{\sum_{j=1}^p E_{\lambda^{k-1}}[\tau_j^2 | \tilde{y}]}}$$

mediante la sustitución de la esperanza condicional con promedios del muestreo de Gibbs [Park y Casella (2008)] donde se sugiere el valor inicial

$$\lambda^{(0)} = \frac{p\sqrt{\widehat{\sigma}_{LS}^2}}{\sum_{j=1}^p |\widehat{\beta}_j^{LS}|}$$

con σ_{LS}^2 y $\widehat{\beta}_j^{LS}$ obtenidos mediante el procedimiento de estimación usual de mínimos cuadrados o en su caso si $p \gg n$ utilizando los estimadores obtenidos mediante la regresión Ridge.

4.2.2. Hiper-aprioris para el parámetro de LASSO

Una alternativa para encontrar explícitamente a λ es dando una hiper-apriori difusa. En Park y Casella (2008) se considera la clase de aprioris Gamma para λ^2 de la forma

$$\pi(\lambda^2) = \frac{\delta^r}{\Gamma(r)} (\lambda^2)^{r-1} \exp(-\delta\lambda^2), \quad \lambda^2 > 0 \quad (r > 0, \delta > 0) \quad (4.4)$$

pues la distribución condicional resultante permite una fácil extensión del muestreador de Gibbs. En caso de utilizar la apriori de escala invariante $\frac{1}{\lambda^2}$ para λ^2 conducirá a una aposteiori inadecuada. Al utilizar la apriori en 4.4 en la jerarquía de 4.3 la distribución condicional completa para λ^2 es Gamma con parámetro de forma $p + r$, parámetro de escala $\sum_{j=1}^p \tau_j^2 / 2 + \delta$ y Esperanza $\frac{p+r}{\sum_{j=1}^p \tau_j^2 / 2 + \delta}$. Con esta especificación λ^2 puede unirse simplemente a otro parámetro en el muestreador de Gibbs, como se presenta en la siguiente sección, ya que la distribución condicional conjunta de los otros parámetros no cambia.

La densidad a priori para λ^2 debe aproximarse a cero lo suficientemente rápido como

4.3. Algoritmo de Gibbs

$\lambda^2 \rightarrow \infty$ (para evitar problemas de convergencia) pero ser relativamente plana y tener alta probabilidad cerca del estimador de máxima verosimilitud [Park y Casella (2008)].

4.3. Algoritmo de Gibbs

EL algoritmo de Gibbs es uno de los métodos más conocidos y utilizados entre las técnicas Cadenas de Markov Monte Carlo (MCMC). En Casella y George (1992) se proporciona una introducción de este método y puede verse con mayor detalle en Casella (2001).

Sea $f(\theta)$ una densidad de la que no se pueden generar valores, completamente conocida, excepto por la constante de integración, pero demasiado compleja como para calcular sus características: marginales, momentos, etc. Estas características pueden abordarse mediante el muestreador de Gibbs, si el vector paramétrico θ puede descomponerse en $k \geq 2$ sub-vectores, $\theta = (\theta_1, \theta_2, \dots, \theta_k)$ de manera que para cada una de las k distribuciones condicionales $f(\theta_i | \theta_1, \dots, \theta_{i-1}, \theta_{i+1}, \dots, \theta_k)$ se pueden generar valores de dichas distribuciones sin excesivo costo computacional. Para esta aplicación, las distribuciones condicionales que se presentan en 4.5 representan a cada una de las $f(\theta_i | \theta_1, \dots, \theta_{i-1}, \theta_{i+1}, \dots, \theta_k)$ con $\theta = (\beta, \sigma^2, \tau, \lambda^2)'$. En la aplicación del muestreador de Gibbs, estas distribuciones condicionales completas son fáciles de muestrear y dependen del vector de respuestas \tilde{y} .

$$\begin{aligned}
 \beta | \sigma^2, \{\tau_j^2\}_{j=1}^p, y &\sim N_p(\tilde{\beta}, \sigma^2 \mathbf{A}^{-1}), \quad \mathbf{A} = X^T X + \mathbf{D}_\tau^{-1}, \quad \tilde{\beta} = \mathbf{A}^{-1} X^T \tilde{y} \\
 \sigma^2 | \beta, \{\tau_j^2\}_{j=1}^p, y &\sim IG((n-1+p)/2, \psi(\beta)/2), \\
 \psi(\beta) &= \|\tilde{y} - X\beta\|^2 + \beta^T \mathbf{D}_\tau^{-1} \beta \\
 \tau_j^{-2} | \beta_j, \sigma^2, \lambda &\sim^{iid} Inv - Gauss(\sqrt{\lambda^2 \sigma^2 / \beta_j^2}, \lambda^2) \\
 \lambda^2 | \tau_1^2, \dots, \tau_p^2 &\sim G(p+r, \sum_{j=1}^p \tau_j^2 / 2 + \delta)
 \end{aligned} \tag{4.5}$$

con de la densidad Gaussiana inversa dada por

$$f(x) = \sqrt{\frac{\lambda}{2\pi}} x^{-3/2} \exp\left\{-\frac{\lambda(x - \mu)^2}{2(\mu)^2 x}\right\}, \quad x > 0$$

Las condicionales completas que se presentan en 4.5 forman la base para una implementación eficiente del muestreador de Gibbs.

4.4. Regresión Lineal Bayesiana

Lasso Bayesiano es visto como un modo posterior en un modelo bayesiano con probabilidad Gaussiana, donde se considera una apriori para β como el producto de densidades independientes doble exponenciales [Park y Casella (2008)]. En contraste la Regresión Lineal Bayesiana (BLR por sus siglas en Inglés) se obtiene asumiendo la misma probabilidad y una apriori para β como el producto de p densidades normales independientes [Cossa *et al.* (2010)].

La diferencia entre las distribuciones tomadas como a priori se presenta en la Figura 4.1. Puede verse que la densidad doble exponencial tiene más masa cerca de cero y colas más gruesas que la distribución Gaussiana lo que produce la seleccion de variables. Desde la perspectiva de BLR, LASSO produce una fuerte contracción en los parámetros de regresión cercanos a cero y menor reducción para aquellos con valores absolutos grandes.

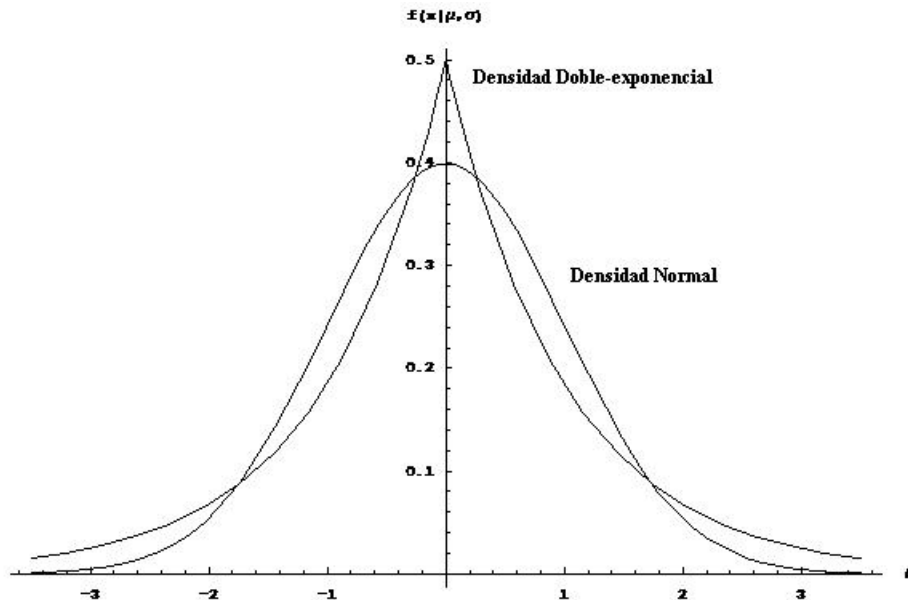


Figura 4.1: Densidad de una distribución normal y doble exponencial (ambas con media cero y varianza igual a uno).

Por facilidad y dado que LASSO Bayesiano ha sido implementado en varios softwares como La Regresión Lineal Bayesiana; BLR [de los Campos y Pérez (2010)] y Blasso [Gramacy (2010)], la metodología utilizada para obtener los $\hat{\beta}_j$ se propone en este trabajo utilizando el paquete BLR, posteriormente aplicar una prueba de permutaciones para probar la significancia de los $\hat{\beta}_j$ obtenidos, por último se utiliza a blasso [Gramacy (2010)] mediante saltos reversibles, que se explica brevemente más adelante, para comparar los valores de los $\hat{\beta}_j$ obtenidos.

4.5. Cadenas de Markov Monte Carlo con Saltos Reversibles

El algoritmo de Cadenas de Markov Monte Carlo con saltos reversibles (RJMCMC), es una generalización del algoritmo Metropolis- Hastings, desarrollado por Green (1995) y consiste en crear una cadena de Markov irreducible, aperiódica que alterna saltos entre distintos modelos con espacios parametrales de diferente dimensión, mientras que satisface la condición de probabilidad de equilibrio, asegurando la convergencia de la distribución límite correcta. Este algoritmo ha sido utilizado ampliamente en la literatura, e.g., Richardson y Green (1997) aplican este algoritmo a mezclas de distribuciones con un número desconocido de componentes. Para esto, implementan dos tipos de movimientos: uno en el que se crea un componente nuevo, o se elimina uno existente, y otro en el que un componente es dividido en dos, o dos componentes son combinados en uno.

El algoritmo RJMCMC permite seleccionar muestras de la distribución de interés definida sobre la unión de subespacios C_r de diferente dimensión n_r . El algoritmo RJMCMC es una generalización de las técnicas MCMC como las encontradas en el problema de selección de variables en regresión o selección de modelos de orden desconocido, en el que la dimensión del espacio parametral puede cambiar de una iteración a otra. Sea M_r , $r = 1, 2, \dots, k_{max}$ los modelos candidatos y θ_r el vector de parámetros desconocido de dimensión n_r del modelo M_r , los que pueden variar de un modelo a otro. Esto es posible cubriendo el espacio de movimientos entre modelos especificando la densidad propuesta, construyendo una función de salto invertible entre los subespacios y calculando la probabilidad de aceptación que satisface la condición de probabilidad de equilibrio. Estas probabilidades se obtienen en esta aplicación mediante la metodología LASSO bayesiano utilizando el paquete BLASSO [Gramacy (2010)]. Existen otras aplicaciones de LASSO Bayesiano que utilizan saltos reversibles, una de ellas puede encontrarse en Chen *et al.* (2011).

Capítulo 5

Prueba de Permutaciones

5.1. Remuestreo

Los métodos de remuestreo nos permiten cuantificar la incertidumbre mediante el cálculo de errores estándar, intervalos de confianza y la realización de pruebas de significación. Estos métodos son útiles pues requieren menos supuestos que los métodos tradicionales y en general dan respuestas más precisas (a veces mucho más exactas). En estadística el remuestreo [Fisher (1935)] es utilizado para estimar la precisión de las estadísticas en la muestra utilizando subconjuntos de los datos disponibles; para la validación de modelos mediante el uso de subconjuntos aleatorios o para hacer pruebas de significancia cambiando las etiquetas de los datos.

Las técnicas comunes de remuestreo incluyen bootstrap, Jackknife, validación cruzada y prueba de permutaciones.

5.2. Prueba de permutaciones

La prueba de permutaciones también llamada prueba de exactitud o prueba de re-aleatorización es un tipo de prueba estadística de significancia en la que se obtiene la distribución del estadístico de prueba bajo la hipótesis nula mediante el cálculo de todos los valores posibles de la prueba estadística en virtud de los reordenamientos de las etiquetas de los puntos de datos observados [Fisher (1935)]. En otras palabras, el método por el cual se asignan los tratamientos a los sujetos en un diseño experimental se refleja en el análisis de ese diseño. Si las etiquetas son intercambiables bajo la hipótesis nula, entonces las pruebas de rendimiento resultantes dan niveles exactos de significación.

5.2. Prueba de permutaciones

Se desea probar que los valores de $\beta_j \neq 0$ no fueron obtenidos al azar, es decir, que al cambiar las etiquetas del vector de respuestas realizando una permutación dentro del vector y , los resultados obtenidos no serán significativos en el valor de $\hat{\beta}$. Para ello se puede utilizar el paquete BLR y llevar a cabo la siguiente prueba de hipótesis mediante la prueba de permutaciones.

En general una prueba de hipótesis es una regla que especifica:

- i) Para qué valores de la muestra no se rechaza H_0 .
- ii) Para qué valores de la muestra se rechaza H_0 en favor de H_a .

Utilizando la probabilidad de cometer un error de tipo I como α_h y la probabilidad de cometer el error de tipo II como β_h , dados por:

$$P(\text{Escoger } \beta_j \neq 0 \text{ dado que } \beta_j = 0) = \alpha_h = 0.05$$

$$P(\text{Escoger } \beta_j = 0 \text{ dado que } \beta_j \neq 0) = \beta_h = 0.95$$

Sea:

$$H_0 : \beta_j = 0$$

$$H_a : \beta_j \neq 0$$

Una vez que se han estimado la distribución de β_j mediante el paquete BLR se puede realizar una permutación dentro del vector de respuestas y para verificar qué indicadores económicos están asociados al GDP. La prueba empírica se puede obtener de las ideas presentadas en Che y Xu (2010) y Churchill y Doerge (1994). Sea el vector $y_{(\text{permutacion}[k])}$ una permutación del vector de respuestas original y , si los indicadores económicos no coinciden con el GDP, los coeficientes de regresión se esperan cerca de cero [Che y Xu (2010)]. Utilizando el vector $y_{(\text{permutacion}[k])}$ se obtiene $\hat{\beta}[k]$ empleando el paquete BLR, utilizando la matriz de covariables X , para obtener el vector de parámetros $\hat{\beta}[k]$, del cual se obtiene el máximo, para nuestros fines llamado Máx $\hat{\beta}[k]$, esto se realiza para cada vector $y_{(\text{permutacion}[k])}$ con $k = 1, \dots, B$, es decir, B veces para generar la distribución empírica.

Utilizando la distribución empírica de la estadística de prueba se obtiene el punto crítico empírico para la prueba dado por \mathbf{Q}_{95} (cuantil del 95 % de la distribución empírica del Max $\hat{\beta}[k]$).

La regla de decisión esta especificada de la siguiente manera: si $|\hat{\beta}_j| \geq \mathbf{Q}_{95}$ se rechaza H_0 , si $|\hat{\beta}_j| \leq \mathbf{Q}_{95}$ no se rechaza H_0 , es decir, el valor absoluto del parámetro $\hat{\beta}_j$ de cada

5.2. Prueba de permutaciones

indicador económico significativo para predecir el GDP está por encima de este punto crítico.

Este procedimiento puede verse en el programa del anexo C.2.

Capítulo 6

Aplicación a Indicadores Económicos

El Producto Interno Bruto, PIB (por sus siglas en Inglés GDP: Gross Domestic Product) se define como el valor de la producción global de un país de bienes y servicios. En esta sección se presenta la aplicación de la metodología Lasso Bayesiano a indicadores económicos relacionados con el GDP de Estados Unidos.

6.1. Definición de variables

Se realizó la búsqueda de variables reales para aplicar la metodología y ajustar el modelo de regresión lineal múltiple en (3.1). La variable respuesta y y las covariables se describen en la Tabla 6.1 y la Tabla 6.2, donde IVA es Inventory Valuation Adjustment y CCA_{adj} es Capital Consumption Adjustment (por sus siglas en Inglés).

Se cuenta con la información de cada variable en periodos trimestrales. Para cada una de ellas se obtuvo la diferencia con respecto al periodo anterior, iniciando el primer trimestre de 2000 hasta el tercer trimestre de 2010 (la primera observación es la diferencia del primer trimestre de 2000 con respecto al segundo trimestre de 2000) con lo cual se obtuvo un total de 42 observaciones disponibles y 63 covariables. Se maneja este periodo dado que la crisis económica y financiera en Estados Unidos se presentó durante esa década.

6.2. Ajuste del modelo

Estos datos presentan el problema en donde existe un mayor número de variables que observaciones, es decir, $p \gg n$, por lo tanto, no es posible obtener la estimación de los parámetros del modelo mediante el método de mínimos cuadrados.

6.2. Ajuste del modelo

Tabla 6.1: Variable respuesta y covariables.

Variable	Definición	Unidad en USA
y	GDP	Billion of \$
x1	Balance on Current Account, NIPA's	Billion of \$
x2	Change in Private Inventories	Billion of \$
x3	Compensation of Employees: Wages and Salary Accruals	Billion of \$
x4	Corporate Business Capital Consumption Adjustment	Billion of \$
x5	Corporate: Consumption of Fixed Capital	Billion of \$
x6	Corporate Inventory Valuation Adjustment	Billion of \$
x7	Corporate Net Cash Flow	Billion of \$
x8	Corporate Profits After Tax	Billion of \$
x9	Corporate Profits After Tax with IVA and CCAdj	Billion of \$
x10	Corporate Profits with IVA and CCAdj	Billion of \$
x11	Disposable Personal Income	Billion of \$
x12	Exports of Goods and Services	Billion of \$
x13	Federal Consumption Expenditures & Gross Investment	Billion of \$
x14	Federal Government: Current Expenditures	Billion of \$
x15	Federal Government Current Receipts	Billion of \$
x16	Federal Grants-in-Aid to State and Local Governments	Billion of \$
x17	Federal National Defense Gross Investment	Billion of \$
x18	Federal Nondefense Gross Investment	Billion of \$
x19	Fixed Private Investment	Billion of \$
x20	Government Consumption Expenditures and Gross Investment	Billion of \$
x21	Government Current Expenditures	Billion of \$
x22	Government Current Receipts	Billion of \$
x23	Real National Defense Gross Investment	Billion of \$
x24	Real Nonresidential Investment: Equipment and Software	Billion of \$
x25	Unemployment rate	Billion of \$
x26	Gross Government Saving	Billion of \$
x27	Gross Private Domestic Investment	Index
x28	Gross Private Domestic Investment: Chain-type Price Index	Billion of \$
x29	Gross Private Saving	Billion of \$
x30	Gross Saving	Billion of \$
x31	National Defense Consumption Expenditures & Gross Investment	Billion of \$
x32	National Income	Billion of \$
x33	National Income: Compensation of Employees, Paid	Billion of \$
x34	Net Corporate Dividends	Billion of \$
x35	Net Federal Government Saving	Billion of \$
x36	Net Government Saving	Billion of \$
x37	Net State and Local Government Saving	Billion of \$
x38	Nondefense Consumption Expenditures & Gross Investment	Billion of \$
x39	Nonfinancial Corporate Business: Profits After Tax	Billion of \$
x40	Personal Consumption Expenditures	Billion of \$

6.3. Aplicando la prueba de permutaciones

Tabla 6.2: Variable respuesta y covariables (Continuación).

Variable	Definición	Unidad en USA
x41	Personal Consumption Expenditures: Chain-type Price Index	Billion of \$
x42	Personal Consumption Expenditures: Chain-type Price Index Less Food and Energy	Billion of \$
x43	Personal Consumption Expenditures: Durable Goods	Billion of \$
x44	Personal Consumption Expenditures: Nondurable Goods	Billion of \$
x45	Personal Consumption Expenditures: Services	Billion of \$
x46	Personal Income	Billion of \$
x47	Personal Saving	Billion of \$
x48	Private Nonresidential Fixed Investment	Billion of \$
x49	Private Residential Fixed Investment	Billion of \$
x50	Proprietors' Income with IVA and CCAdj	Billion of \$
x51	Real Change in Private Inventories	Billion of \$
x52	Real Disposable Personal Income	Billion of \$
x53	Real Exports of Goods & Services	Billion of \$
x54	Real Federal Consumption Expenditures & Gross Investment	Billion of \$
x55	Real Federal Nondefense Gross Investment	Billion of \$
x56	Real Final Sales of Domestic Product	Billion of \$
x57	Real Government Consumption Expenditures & Gross Investment	Billion of \$
x58	Real Gross Domestic Product	Billion of \$
x59	Real Gross Private Domestic Investment	Billion of \$
x60	State & Local Consumption Expenditures & Gross Investment	Billion of \$
x61	State & Local Government Current Expenditures	Billion of \$
x62	State & Local Government Gross Investment	Billion of \$
x63	Rental Income of Persons with CCAdj	Billion of \$

Se sabe que el modelo utilizado en el paquete BLR es un modelo generalizado que incluye a LASSO Bayesiano con la diferencia de que han utilizado diferentes a priori para los β_j [Cossa *et al.* (2010)], por esta razón, BLR fue utilizado para ajustar el modelo. Los resultados obtenidos con BLR en R se presentan en la Figura 6.1

Al ajustar el modelo en (3.1) con la metodología LASSO Bayesiano (BLR), con $\lambda = 72.70293$, se obtuvieron las medias y desviaciones estándar de cada uno de los β_j con lo cual se realizó el diagrama la Figura 6.2.

Para elegir las variables que se incluyan en el modelo se prosiguió a implementar la prueba de permutaciones y verificar la significancia de los $\hat{\beta}_j$ que se observan en la Figura 6.1

6.3. Aplicando la prueba de permutaciones

Al aplicar la prueba de permutaciones a los datos se obtuvo la distribución del estadístico de prueba (con $B = 5000$) cuya distribución se muestra en la Figura 6.3, utilizando esta distribución se obtiene $\mathbf{Q}_{95} = 0.3698551$.

6.3. Aplicando la prueba de permutaciones

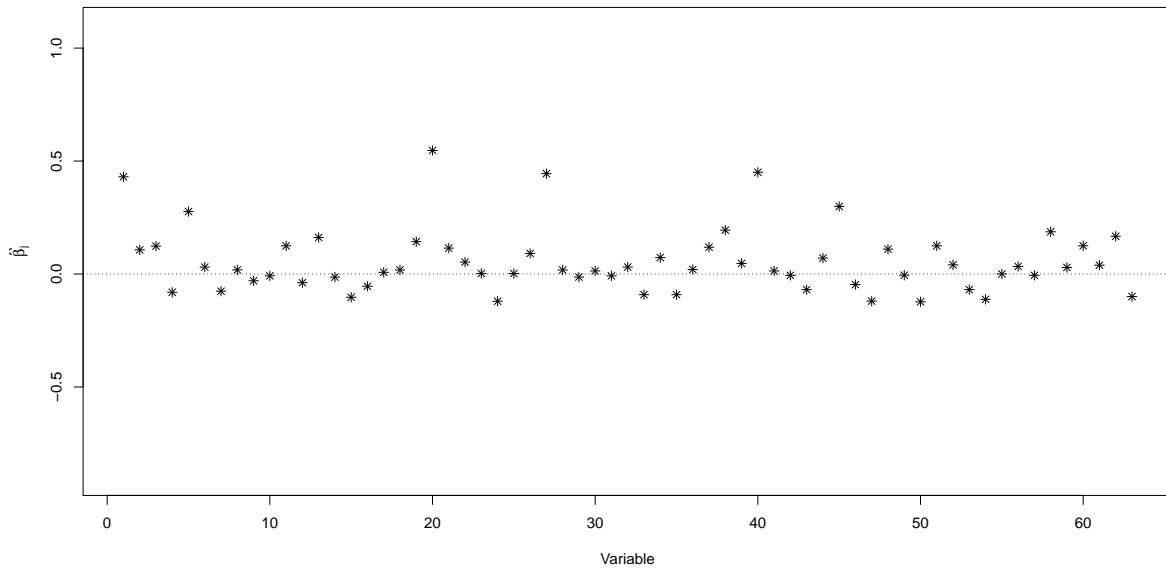


Figura 6.1: Efecto estimado de las covariables en el modelo $y = X\beta + \epsilon$ obtenidos utilizando LASSO Bayesiano.

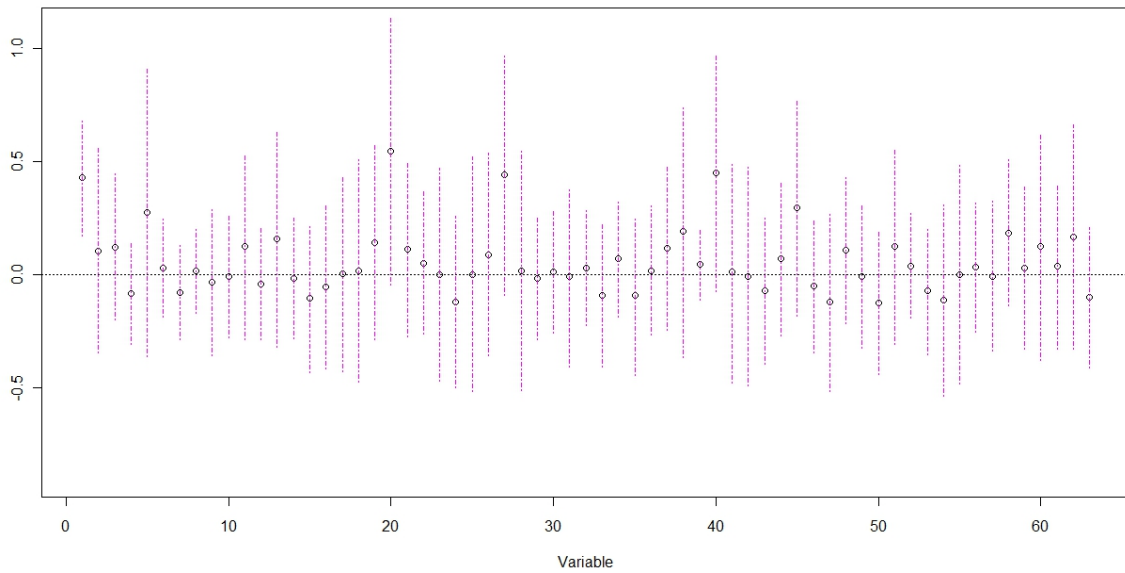


Figura 6.2: Media de $\hat{\beta}_j$, cada uno con cuantil del 2.5% al 97.5% representado por una línea punteada.

Al obtener el cuantil del 95% de los datos el valor obtenido fue $Q_{95} = 0.3698551$, éste se observa en la Figura 6.3 con una línea vertical.

6.3. Aplicando la prueba de permutaciones

Este cuantil nos sirve para valorar a cada uno de los $\hat{\beta}_j$ obtenidos en el ajuste original y obtener los $\hat{\beta}_j$ significativos en el modelo, es decir, los valores de $|\hat{\beta}_j|$ mayores al cuantil del 95 % serán significativos en el modelo.

Los valores de $\hat{\beta}_j$ que están por encima de $\mathbf{Q}_{95} = 0.3698551$ se presentan en la Tabla 6.3.

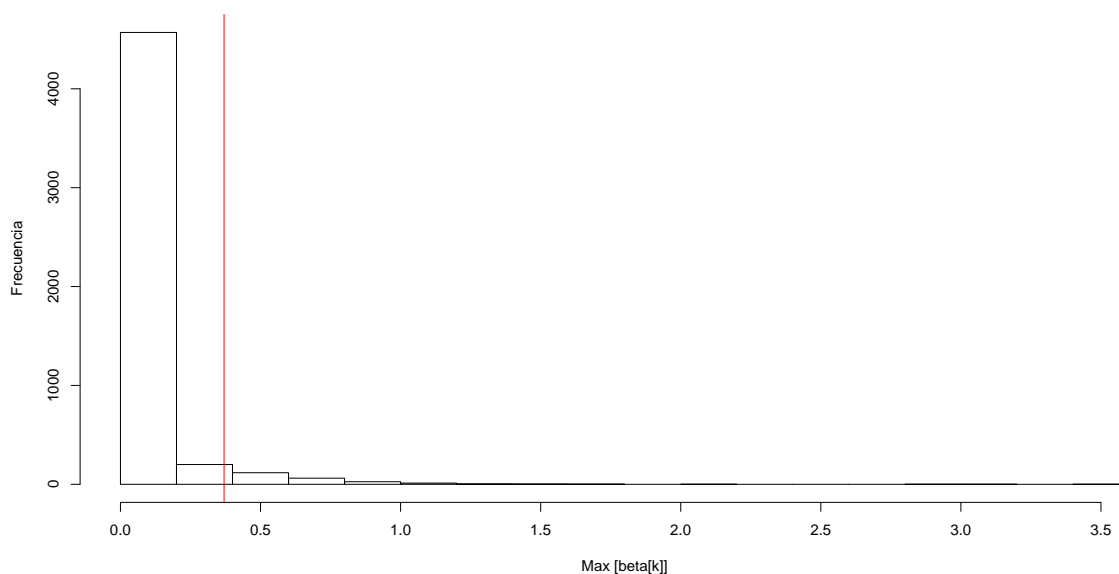


Figura 6.3: Distribución empírica de la estadística de prueba.

Tabla 6.3: Valores de $|\hat{\beta}_j| > \mathbf{Q}_{95}$ obtenidos para el modelo LASSO Bayesiano

Parámetro	β_1	β_{20}	β_{27}	β_{40}
Valor	0.4302071	0.5468196	0.4443131	0.4502355

Los nombres de las variables con $\hat{\beta}_j > \mathbf{Q}_{95} = 0.3698551$ se presentan en la Tabla 6.4.

Tabla 6.4: Variables con parámetro $|\hat{\beta}_j| > \mathbf{Q}_{95}$

Parámetro	variable	Definición
β_1	x_1	Balance on Current Account, NIPA's
β_{20}	x_{20}	Government Consumption Expenditures and Gross Investment
β_{27}	x_{27}	Gross Private Domestic Investment
β_{40}	x_{40}	Personal Consumption Expenditures

6.4. Blasso mediante saltos reversibles

Para verificar el ajuste del modelo se realizó el ajuste utilizando el paquete Blasso [Gramacy (2010)] mediante saltos reversibles que se comentan en la sección 4.5. Con este paquete se pueden verificar los resultados obtenidos mediante LASSO Bayesiano de la sección 6.2.

Los resultados de las medias obtenidas para cada β_j se presentan en la Figura 6.4. Se presenta también el diagrama de cajas en la Figura 6.5.

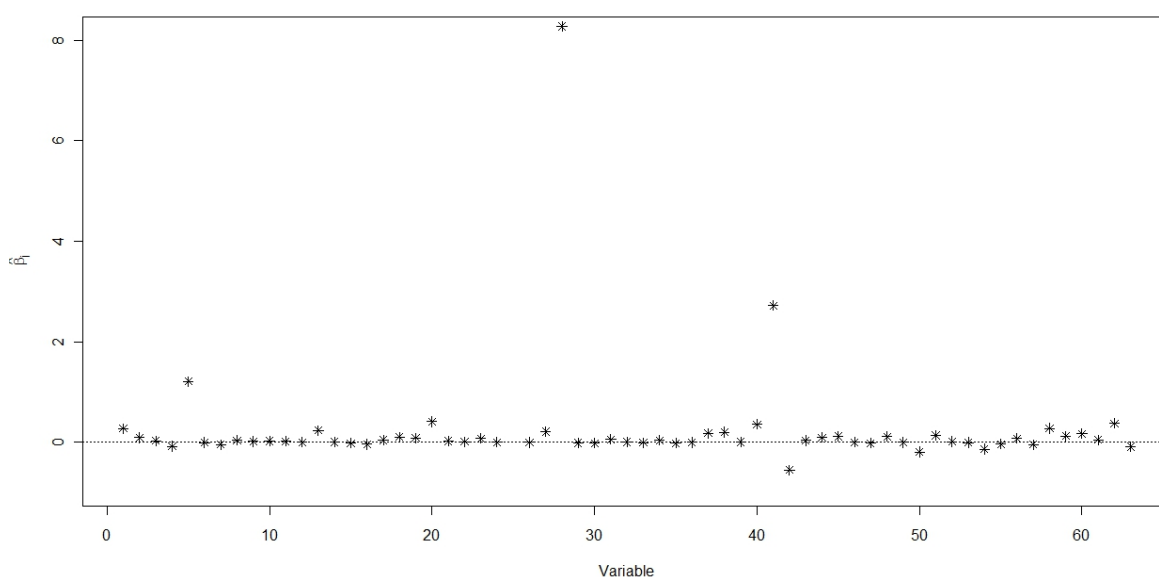


Figura 6.4: Coeficientes de Lasso Bayesiano utilizando el paquete Blasso.

La probabilidad de que $\beta_j \neq 0$ es un dato importante que se obtiene del paquete Blasso, ya que da gran aporte para tomar decisiones sobre los coeficientes $\hat{\beta}_j$, esto es, no importa si el valor de $\hat{\beta}_j$ es pequeño si su probabilidad de ser diferente de cero es relativamente alta, es aquí donde cada analista puede tomar la decisión de qué β_j debe incluirse en el modelo. Para este trabajo se tomó un punto crítico para el valor de la $P(\beta_j \neq 0)$ y poder seleccionar los β_j que se incluyan en el modelo, es decir, se tomarán los β_j con $P(\beta_j \neq 0) > 0.7$.

En la Figura 6.4 se presenta la media de $\beta_{28} = 8.278$, aunque es un valor muy alto, no se tomó este parámetro como parte del modelo dado que $P(\beta_{28} \neq 0) = 0.662 < 0.7$, lo mismo sucedió para $\beta_{41} = 2.722$, dado que $P(\beta_{28} \neq 0) = 0.441 < 0.7$.

La probabilidad de que $\beta_j \neq 0$ se presenta en la Figura 6.6, los parámetros que tienen probabilidad de ser diferentes de cero mayor a 0.7 se presentan en la Tabla 6.5.

6.4. Blasso mediante saltos reversibles

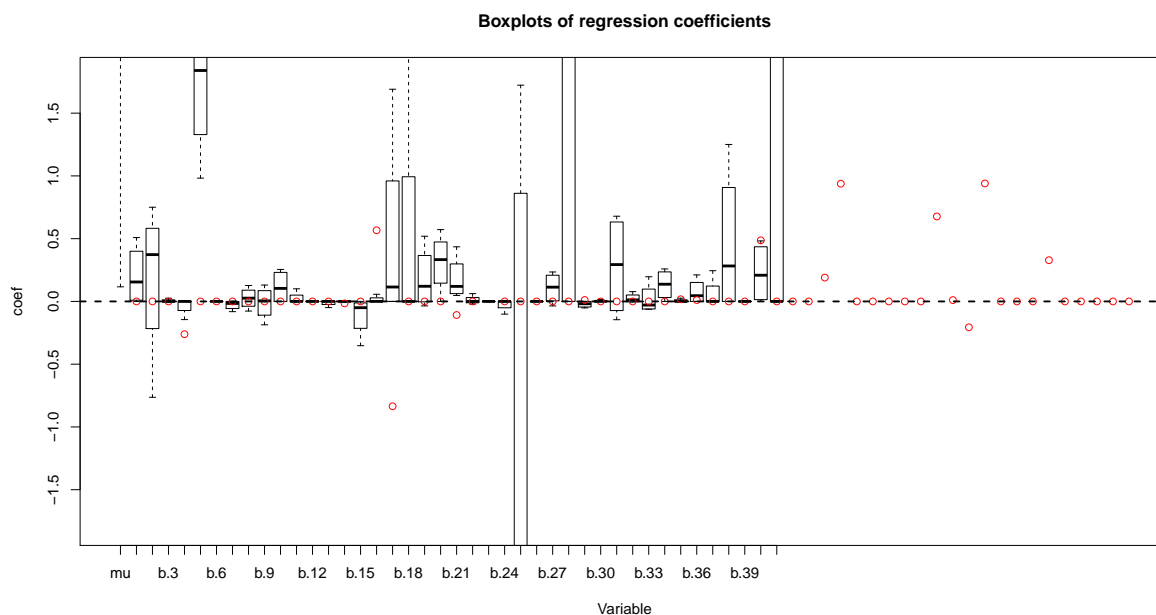


Figura 6.5: Diagrama de cajas para β_j en el modelo Blasso.

Tabla 6.5: Probabilidad de $\beta_j \neq 0$.

Parámetro	β_1	β_5	β_{20}	β_{27}	β_{40}	β_{58}
Probabilidad	0.904	0.807	0.753	0.778	0.881	0.874

Los nombres de las variables con con $P(\beta_j \neq 0) > 0.7$ se muestran en la Tabla 6.6

Tabla 6.6: Variables con parámetro $\beta_j \neq 0$

Parámetro	variable	Definición
β_1	x_1	Balance on Current Account, NIPA's
β_5	x_5	Corporate: Consumption of Fixed Capital
β_{20}	x_{20}	Government Consumption Expenditures and Gross Investment
β_{27}	x_{27}	Gross Private Domestic Investment
β_{40}	x_{40}	Personal Consumption Expenditures
β_{58}	x_{58}	Real Gross Domestic Product

Los resultados estadísticos de la estimación de los $\hat{\beta}_j$ se presentan en la Tabla 6.7.

Como puede verse estos valores de $\hat{\beta}_j$ son similares a los obtenidos con la metodología de Lasso Bayesiano aplicando el paquete BLR que se presentan en la Tabla 6.3 y 6.4.

Otra forma de ver estos resultados es mediante una gráfica que compare los valores observados contra los valores predichos puede usarse para comparar las distribuciones inferidas directamente de dos conjuntos de observaciones, es decir, funciona para comparar el va-

6.4. Blasso mediante saltos reversibles

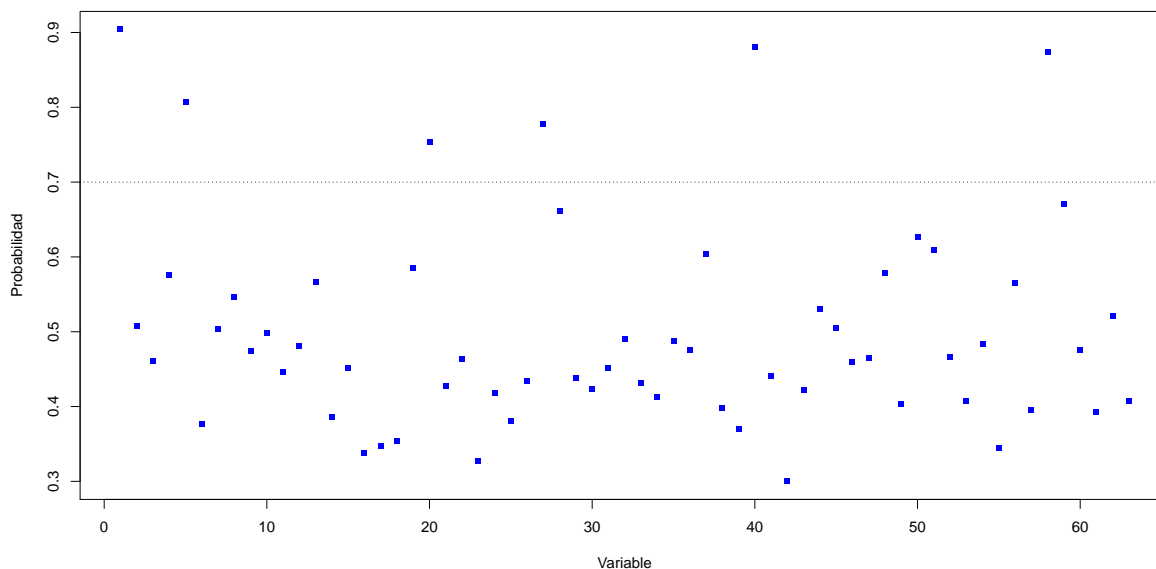


Figura 6.6: Probabilidad de $\beta_j \neq 0$.

Tabla 6.7: Información de los parámetros con $P(\beta_j \neq 0) > 0.7$ obtenidos en R utilizando el paquete Blasso

Estadística parámetro	β_1	β_5	β_{20}	β_{27}	β_{40}	β_{58}
Min.	-0.1337	-0.7628	-0.7075	-0.1069	-0.1925	-0.1798
1st Qu.	0.1598	0.2641	0.0000	0.0000	0.1734	0.1249
Median	0.2874	1.2215	0.3811	0.1690	0.3814	0.2921
Mean	0.2732	1.2098	0.4120	0.2140	0.3601	0.2788
3rd Qu.	0.3857	1.9506	0.6952	0.3682	0.5360	0.4118
Max.	0.8059	4.1957	1.7901	1.0343	1.0961	0.9614

lor de y con \hat{y} , la gráfica de valores observados contra valores predichos para el ajuste con BLR se presenta en la Figura 6.7, por comparación se presentan las graficas para el ajuste utilizando la metodología LASSO clásico y regresión Ridge en la Figura 6.8 y la Figura 6.9. Se presenta también la gráfica de Blasso mediante saltos reversibles en la Figura 6.10.

Al ajustar el modelo con la metodología LASSO Bayesiano, LASSO clásico, regresión Ridge y Blasso [Anexo C.1], se obtuvo la correlación (y, \hat{y}) como una medida del ajuste del modelo, los resultados se presentan en la Tabla 6.8.

Tabla 6.8: Correlación entre y y \hat{y} .

	LASSO Bayesiano	LASSO Clásico	Ridge	Blasso
Correlación (y, \hat{y})	0.9972711	0.997193	0.9970947	0.996835

6.4. Blasso mediante saltos reversibles

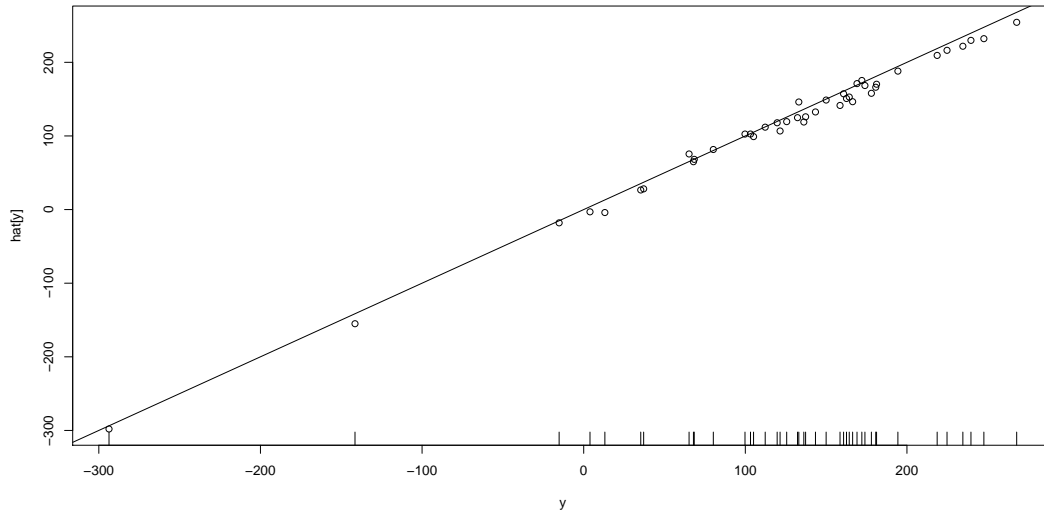


Figura 6.7: Grafica de y vs \hat{y} utilizando la metodología LASSO Bayesiano.

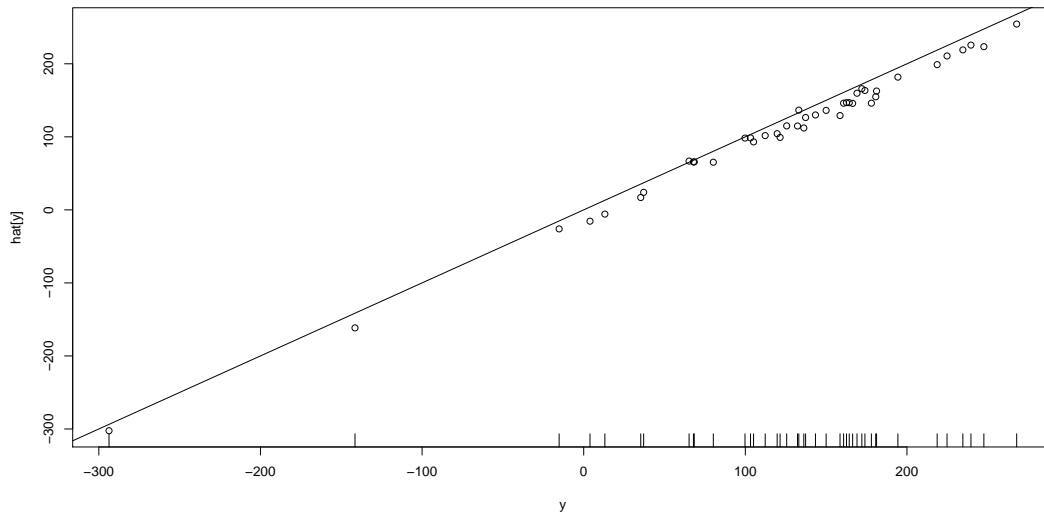


Figura 6.8: Grafica de y vs \hat{y} utilizando la metodología LASSO Clásico.

La mejor correlación se ha obtenido mediante la metodología LASSO Bayesiano utilizando el paquete BLR.

6.4. Blasso mediante saltos reversibles

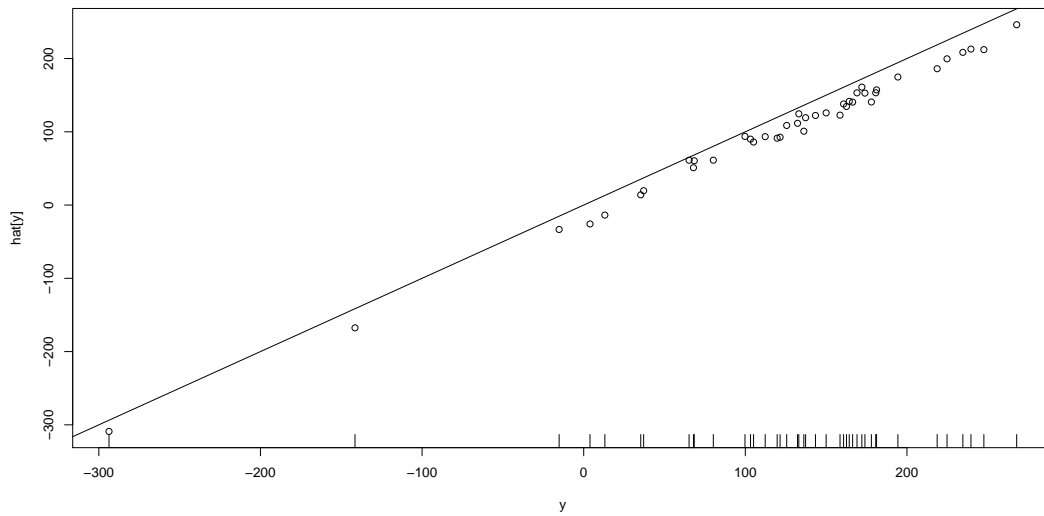


Figura 6.9: Grafica de y vs \hat{y} utilizando la regresión Ridge.

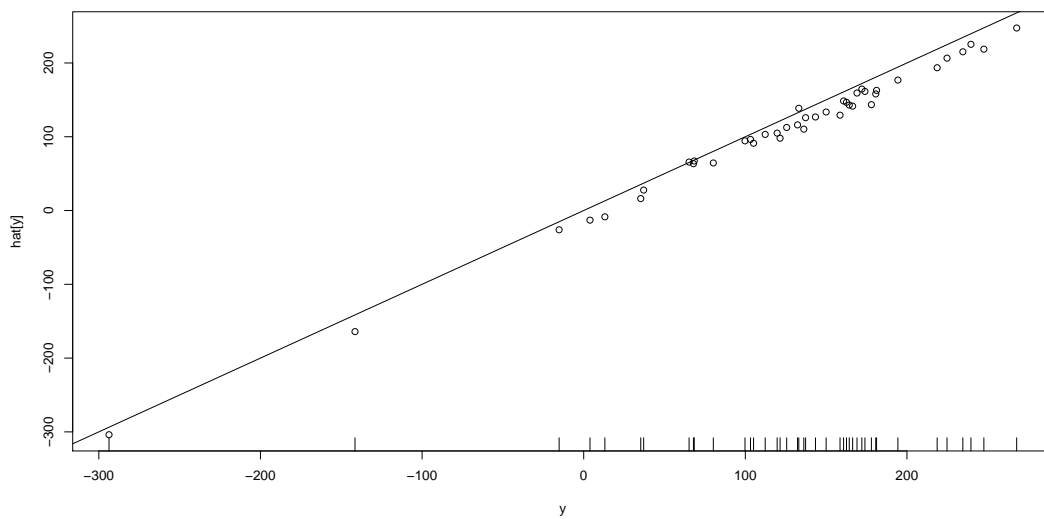


Figura 6.10: Grafica de y vs \hat{y} utilizando el paquete Blasso.

Capítulo 7

Discusión de Resultados

Las variables resultantes para el modelo de regresión se presentan en la Tabla 7.1 . Es importante mencionar que se ha optado por estas variables ya que han resultado del ajuste de ambos paquetes aplicados, BLASSO y BLR. Para discutir el resultado obtenido se da a continuación una breve descripción de cada una de estas variables:

Tabla 7.1: Variables finales obtenidas en el ajuste del modelo

Parámetro	variable	Definición
β_1	x_1	Balance on Current Account, NIPA's
β_{20}	x_{20}	Government Consumption Expenditures and Gross Investment
β_{27}	x_{27}	Gross Private Domestic Investment
β_{40}	x_{40}	Personal Consumption Expenditures

x_1 -Balance on Current Account, NIPA's : Balance de cuentas corrientes del NIPA (National income and Product Accounts por sus siglas en Inglés). Este balance muestra el resultado de la composición de la producción y la distribución de los ingresos obtenidos en la producción. Por lo tanto, representan un elemento crítico de las cuentas económicas de EE.UU., está diseñado para proporcionar una imagen coherente y global de la economía de la nación.

x_{20} -Government Current Expenditures: Gastos corrientes de gobierno, son aquellos gastos que realiza el gobierno día con día, es decir, el gasto en artículos recurrentes, incluyendo sueldos y salarios recurrentes, gastos en consumibles y artículos de uso diario en la prestación de bienes y servicios por el gobierno.

x_{27} -Gross private domestic investment: Inversión interna bruta privada, es la medida de la inversión utilizada para calcular el PIB. Este es un componente importante del PIB, ya que proporciona un indicador de la capacidad productiva futura de la economía. Incluye las compras de reemplazo, más las adiciones netas de bienes de capital, más las inversiones en inventarios. Por lo general, equivale a entre 15 y 18 por ciento del PIB. La inversión

7. Discusión de Resultados

neta es la inversión bruta menos la depreciación. La inversión interna bruta privada incluye tres tipos de inversión; Inversiones no residenciales: Gastos de las empresas en máquinas, herramientas, etc.; Inversión en viviendas: Incluye los gastos de los hogares y las empresas en apartamentos, edificios y fábricas nuevas; Cambio en los inventarios: El cambio en los inventarios de la empresa en un período determinado. (Inventario: es la mercancía que es producida por las empresas, pero los productos que se mantienen para su venta posterior).

x_{40} -Personal Consumption Expenditures: Gastos de Consumo Personal. Es una medida de las variaciones de precios de bienes y servicios de consumo. Los gastos de consumo personal consistirá en los gastos reales y ficticios de los hogares, la medida incluye los datos relativos a bienes duraderos, no duraderos y servicios. Es esencialmente una medida de bienes y servicios dirigidos a personas y consumidos por los individuos.

Aunque en la aplicación de la metodología LASSO Bayesiano aparecen 2 variables más en la salida de los resultados y estas variables no aparecen en la aplicación de BLR, es necesario definir las y dejar abierta la posibilidad de uso de ellas, estas són:

x_5 -Consumption of Fixed Capital: El consumo de capital fijo es la declinación, durante el transcurso del período contable, en el valor actual de acciones de activos fijos que posee y utiliza un productor como consecuencia del deterioro físico, obsolescencia normal o de daños accidentales normales.

x_{58} -Real Gross Domestic Product: El producto Interno bruto real es una medida macroeconómica del valor de la economía de la producción ajustada por los cambios de precios (es decir, la inflación o la deflación).

Es claro notar que el Producto Interno Bruto Real es directamente proporcional al Producto interno bruto, por lo mismo no tiene sentido utilizar esta variable dentro del ajuste del modelo, es preferible quitar esta variable del modelo para beneficio del ajuste dado que se puede obtener mediante una simple formula dada: $PIB = PIB_{real} \times Precio$.

Para la variable: Consumo de Capital Fijo, dado que los activos fijos no entran dentro de los bienes y servicios producidos en un país y por la formula $PIB - D = PIN$ donde D es la depreciación activos fijos y PIN es el Producto Interno Neto. Es lógico que la variable pueda ser descartada del modelo puesto que no forma parte del PIB de interés en este trabajo.

Los resultados obtenidos mediante el ajuste del modelo LASSO Bayesiano tanto en BLASSO como BLR se tomarán en este trabajo con los $\hat{\beta}_j$ de la Tabla 7.2, en donde puede observarse el valor de la mediana obtenida con BLASSO y el valor del coeficiente obtenido con BLR.

Con las definiciones anteriores es fácil ver la relación entre x_1 , x_{20} , x_{27} , x_{40} y el PIB de E.E. U.U. Es claro que ahora podemos identificar un modelo factible para el PIB y el

7. Discusión de Resultados

Tabla 7.2: Media (BLASSO) y Coeficiente del parámetro (BLR) obtenidas en el ajuste del modelo

Variable	Media (Blasso)	Coeficiente(BLR)
Balance on Current Account, NIPA's	0.2732	0.4302
Government Consumption Expenditures and Gross Investment	0.4120	0.5468
Gross Private Domestic Investment	0.2140	0.4443
Personal Consumption Expenditures	0.3601	0.4502

conjunto de variables definidas arriba. Aunque la opción queda abierta para definir otras posibilidades dentro del modelo, por ejemplo, si se realiza el ajuste de este modelo con un subconjunto específico de interés tomado de las 63 variables disponibles en este trabajo, puede hallarse un modelo específico interesante.

Capítulo 8

Conclusiones y Recomendaciones

De los resultados obtenidos en esta investigación se puede concluir lo siguiente:

- Se encontró que las variables que tienen mayor influencia en el modelo que tiene como variable respuesta al PIB son: Balance de cuentas corrientes del NIPA, Gastos corrientes de gobierno, Inversión interna bruta privada y Gastos de Consumo Personal.
- Algunas de estas variables coinciden con lo citado en Baumohl (2008), donde se menciona como el crecimiento de los Gastos de consumo personal están relacionados con la mejora de la economía, o como el PIB es una de las variables más sensibles al cambio de la inversión bruta privada.
- Se obtuvo un buen ajuste del modelo lineal para el PIB y las variables presentadas en la Tabla 7.2 ya que la correlación del ajuste medida entre y y \hat{y} utilizando el modelo generalizado que incluye a LASSO bayesiano, BLR es de 0.9972711.
- El problema donde $p \gg n$ puede abordarse de forma sencilla aplicando la metodología LASSO Bayesiano mediante el paquete Regresión bayesiana lineal para llegar a un buen ajuste con poco trabajo computacional y en poco tiempo.
- Es posible utilizar $P(\beta_j \neq 0)$ obtenida en la salida del ajuste del modelo con la metodología LASSO bayesiano y el paquete BLASSO para los regresores $\hat{\beta}_j$ como un parámetro que puede ayudar a la toma de decisiones sobre las variables que se deben incluir en el modelo final, vea la Tabla 6.5.
- Existe la posibilidad de generar nuevos modelos para el caso presentado en este escrito utilizando subconjuntos de las variables descritas en este trabajo.
- Utilizar ajustes de modelos mediante el uso de métodos bayesianos es una gran ventaja computacional al utilizar el muestreador de Gibbs.

Referencias

- Baumohl, B. (2008). *The Secrets of Economic Indicators*. Wharton School Publishing, London, primera edición.
- Casella, G. (2001). Empirical Bayes Gibbs Sampling. *Biostatistics*, 2, 4, 485–500.
- Casella, G. y George, E. I. (1992). Explaining the Gibbs Sampler. *The American Statistician*, 46, 167–174.
- Che, X. y Xu, S. (2010). Significance Test and Genome Selection in Bayesian Shrinkage Analysis. *International Journal of Plant Genomics*, doi:10.1155/2010/893206.
- Chen, X., Wang, Z. y McKeown, M. J. (2011). A Bayesian Lasso via Reversible-Jump MCMC. *Elsevier*, doi:10.1016/j.sigpro.2011.02.014.
- Churchill, G. A. y Doerge, R. W. (1994). Empirical threshold values for quantitative trait mapping. *Genetics*, 138, 963–971.
- Crossa, J., de los Campos, G., Pérez, P., Gianola, D., Atlin, G., Burgueño, J., Araus, J. L. y Makumbi, D. (2010). Prediction of Genetic Values of Quantitative Traits in Plant Breeding Using Pedigree Molecular Markers. *Genetics*, 186, 2, 713–724.
- de los Campos, G., Naya, H., Gianola, D., Crossa, J., Legarra, A., Manfredi, E., Weigel, K. y Cotes, J. M. (2009). Predicting Quantitative Traits With Regression Models for Dense Molecular Markers and Pedigree. *Genetics*, DOI: 10.1534/genetics.109.101501.
- de los Campos, G. y Pérez, P. (2010). *R: Package ‘BLR’: Bayesian Linear Regression*. R Foundation for Statistical Computing, Vienna, Austria.
- Efron, B., Hastie, T., Johnstone, I. y Tibshirani, R. (2004). Least Angle Regression. *The Annals of Statistics*, 32, 2, 407–499.
- Fisher, R. A. (1935). *The Design of Experiments*. Hafner Press, New York.
- Gramacy, R. B. (2010). *R: Package ‘monomvn’, Blasso*.
- Gramacy, R. B., Lee, J. H. y Silva, R. (2009). On Estimating Covariances Between Many Assets with Histories of Highly Variable Length. *University of Cambridge, Investment London, University College London*.

Referencias

- Gramacy, R. B. y Pantaleo, E. (2010). Shrinkage Regression for Multivariate Inference with Missing Data, and an Application to Portfolio Balancing. *Bayesian Analysis*, 5 Number 2, 237–262.
- Green, P. (1995). Reversible Jump Markov Chain Monte Carlo Computation and Bayesian Model Determination. *Biometrics*, 82, 4, 711–732.
- Hoerl, A. E. (1964). Ridge Analysis. *Chemical Engineering progress*, 60, 67–77.
- Hoerl, A. E. y Kennard, R. W. (1970). Ridge Regression: Biased Estimation for Nonorthogonal Problems. *Technometrics*, 12, 55–67.
- Hotelling, H. (1933). Analysis of a Complex of Statistical Variables into Principal Components. *Journal of Educational Psychology*, 24, 6, 417–441.
- Lee, P. M. (2004). *Bayesian Statistics An Introduction*, tomo 2. Hodder Arnold, New York, segunda edición.
- Marquardt, D. W. y Snee, R. S. (1975). Ridge Regression in Practice. *American Statistician*, 29, 1, 3–20.
- Park, T. y Casella, G. (2008). The Bayesian Lasso. *Journal of the American Statistical Association*, 103, 482, 681–686.
- R Development Core Team (2011). *R: A Language and Environment for Statistical Computing*. R foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0.
- Richardson, S. y Green, P. (1997). On Bayesian Analysis of Mixtures with an Unknown Number of Components (with discussion). *Journal of the Royal Statistical Society*, 59, 731–792.
- Tibshirani, R. (1996). Regression Shrinkage and Selection via the Lasso. *Journal of the Royal Statistical Society*, 58, 1, 267–288.
- Wold, H. (1975). Path Model with Latent Variables: The NIPAS approach. *Quantitative Sociology: International Perspectives on Mathematical and Statistical Modeling*, 307–357.

Anexos

Anexo A: Distribución Doble Exponencial

La distribución de Laplace resulta de la diferencia de dos variables exponenciales aleatorias, independientes e idénticamente distribuidas.

Función de densidad

$$f_X(x|\theta, \sigma) = (2\sigma)^{-1} \exp\left(-\frac{|x - \theta|}{\sigma}\right), \theta > 0.$$

Función de distribución

$$F_X(x) = \begin{cases} \frac{1}{2} \exp[-(\theta - x)/\sigma], & x \leq \theta \\ 1 - \frac{1}{2} \exp[-(x - \theta)/\sigma], & x \geq \theta \end{cases}$$

Función generatriz de momentos

$$m_X(t) = \frac{\exp(\theta t)}{1 - \sigma^2 t^2}, |t| < \frac{1}{\sigma}$$

Media

$$E\{|X|\} = \theta, \text{ Desviación media}$$

Mediana

$$\text{Mediana} = \theta$$

Varianza

$$E\{(X - \mu)^2\} = 2\sigma^2$$

Anexo B: Actualización en Bayes Empírico

El método Monte carlo EM para la estimación en bayes empírico de los hiperparámetros propuesto por Casella (2001) trata esencialmente a los parámetros como datos faltantes posteriormente utiliza iterativamente el algoritmo EM para aproximar los hiperparámetros sustituyendo los estimadores Monte Carlo que no pueden ser calculados explícitamente por algún valor esperado. El muestreador de Gibbs en el LASSO Bayesiano es utilizado para estimar estos valores esperados. La jerarquía de sección 4 en conjunto con la apriori Gamma inversa

$$\pi(\sigma^2) = \frac{\gamma^a}{\Gamma(a)} (\sigma^2)^{-a-1} \exp(-\frac{\gamma}{\sigma^2}), \quad \sigma^2, \quad a > 0, \quad \gamma > 0$$

producen los datos completos en la Log-verosimilitud

$$\begin{aligned} & -((n + p - 1)/2 + a + 1) \ln(\sigma^2) - \frac{1}{\sigma^2} (\|\tilde{y} - X\beta\|_2^2/2 + \gamma) \\ & - \frac{1}{2} \sum_{j=1}^p \ln(\tau_j^2) - \frac{1}{2} \sum_{j=1}^p \frac{\beta_j^2}{\sigma^2 \tau_j^2} + p \ln(\lambda^2) - \frac{\lambda^2}{2} \sum_{j=1}^p \tau_j^2 \end{aligned}$$

La a priori de escala invariante $1/\sigma^2$, es obtenida formalmente tomando $a = 0$ y $\gamma = 0$ puede ser utilizada alternativamente sin invalidar el siguiente tratamiento.

Al realizar la iteración k se utiliza $\lambda^{(k-1)}$ de la iteración previa (o el valor inicial, si $k = 1$) Idealmente el Paso-E implica tomar el valor esperado de la Log-verosimilitud, condicional sobre \tilde{y} y bajo $\lambda^{(k-1)}$ para obtener

$$Q(\lambda | \lambda^{(k-1)}) = p \ln(\lambda^2) - \frac{\lambda^2}{2} \sum_{j=1}^p E_{\lambda^{(k-1)}}[\tau_j^2 | \tilde{y}] + \text{terminos que no dependen de } \lambda$$

El paso-M maximiza esta expresión sobre λ para producir la siguiente estimación $\lambda^{(k)}$. Para este caso existe una simple solución analítica:

$$\lambda^{(k)} = \sqrt{\frac{2p}{\sum_{j=1}^p E_{\lambda^{k-1}}[\tau_j^2 | \tilde{y}]}}$$

Obviamente, las esperanzas condicionales son justamente las esperanzas posteriores bajo el hiperparámetro $\lambda^{(k-1)}$ y por lo tanto pueden estimarse mediante promedios muestrales de una corrida del muestreador de Gibbs de la sección 4.

Anexo C: Rutinas en R-2.12.1

C.1. Rutinas en R-2.12.1 para implementar las metodologías LASSO, LASSO Bayesiano, Regresión Ridge y BLR

Se desarrollaron algunas rutinas en el lenguaje R para correr las metodologías LASSO, LASSO Bayesiano, regresión Ridge y BLR así como para realizar algunos otros cálculos.

Los paquetes que se deben cargar en R son: regress, MASS, Lars, pls, monomvn.

A continuación se presentan las rutinas en R.

```
//Programa para correr la metodolgia LASSO
reg.las <- regress(X, y, method="lasso")
reg.las$lambda
reg.las$b
plot(reg.las$b)
```

```
// Programa para correr la metodología LASSO Bayesiano
reg.blas<-blasso(X,y)
plot(reg.blas,burnin=200,which="coef")
s<-summary(reg.blas)
s$coef
s$bn0
plot(reg.blas, burnin=200, which="m")
```

```
//Programa para correr la metodología de la regresión Ridge
reg.rid <- regress(X, y, method="ridge")
reg.rid$b
plot(reg.rid$b)
```

```
//Programa para correr la metodología en BLR
fm<-BLR(y,XL=X, nIter=20000)
fm$bL
fm$mu
plot(fm$bL, col=4, pch=9)
abline(h=0, lty=3)
fm$SD.bL
```

C.2. Rutina en R-2.12.1 para realizar la Prueba de Permutaciones

```
//Programa para la prueba de permutaciones
fm<-BLR(y,XL=X, nIter=10000)

B=5000
outBL=matrix(NA,nrow=B,ncol=length(fm$bL))
n=length(y)
for(i in 1:B)
{
  index=sample(1:n)
  y_permuta=y[index]
  fm_permuta<-BLR(y_permuta,XL=X, nIter=5000)
  outBL[i,]=(fm_permuta$bL/fm_permuta$SD.bL)^2
}
apply(outBL,1,max)
hist(apply(outBL,1,max), main="Histograma de Prueba de Permutaciones con 5000 corridas")
abline(v=quantile(apply(outBL,1,max), 0.95), col=2)
quantile(apply(outBL,1,max), 0.95)
fm$bL
plot(fm$bL, col=4, pch=8)
abline(h=quantile(apply(outBL,1,max), 0.95), lty=3)
which(abs(fm$bL)>quantile(apply(outBL,1,max), 0.95))
```