

Capítulo 4

Diseños dinámicos

Del mismo modo que en el capítulo 3 hemos intentado explicar qué valor tienen los diseños factoriales en cuanto a su aportación “estática” al avance del conocimiento secuencial en el problema de las superficies de respuesta, aquí proponemos un primer “mapa” conceptual de cuáles pueden ser algunos puntos de partida para entender la forma de definir estrategias de diseño para estos problemas. El desarrollo del presente creemos que servirá para ver con mejor claridad la definición del problema de estudio —el encadenamiento de diseños, que hemos llamado enfoque dinámico— para luego proponer y estudiar estrategias concretas de exploración secuencial de las superficies, tema que abordaremos en el capítulo siguiente.

4.1. Introducción

Habiendo presentado los principios básicos tanto de la *MSR* (en su enfoque clásico para modelos normales), como así también de los *MLG* (y en particular, los detalles más salientes de los *MDB*), intentaremos esbozar en este capítulo un primer mapa conceptual que describa de forma cualitativa cuáles serán los pasos a seguir para enlazar la *MSR* desde los *MDB*.

En éste, pondremos el énfasis en el encadenamiento lógico de conceptos orientándolo a nuestros objetivos: desde los diseños de primer orden hasta los de segundo, tanto para experimentos que tengan uno como dos factores de variabilidad. Y como cierre del capítulo, daremos una idea también tipo “mapa conceptual” acerca de cómo se pueden combinar convenientemente estos enfoques estáticos y dar lugar a otro dinámico para el estudio de superficies de respuesta. Este encadenamiento de enfoques estáticos constituirá el mecanismo de análisis que se utiliza en *MSR*, cuanto menos, desde un punto de vista “clásico”.

En el capítulo siguiente, abordaremos con más profundidad los fundamentos teóri-

cos más relevantes sobre esta metodología secuencial, proyectándola especialmente sobre los *MDB* y utilizando el enfoque de estudio que proporcionan los *MLG*. Mediante un proceso de simulación de datos, prepararemos el terreno para estudiar alternativas de estrategias de estudio e intentar justificar la conveniencia de la utilización de las mismas.

4.2. Motivación

En los casos de estudios de superficies de respuesta para datos con distribución normal, existen varios recursos —en cuanto a disponibilidad de datos concretos— para ensayar estrategias diferentes de exploración secuencial de superficies de respuesta. Basta solamente con consultar la bibliografía clásica sobre este tema para encontrar numerosos ejemplos de aplicaciones prácticas, en disciplinas tanto industriales y biológicas como de varias ciencias sociales¹.

Pero para el caso de los *MDB*, no hemos podido encontrar —al menos, hasta la entrega final de este trabajo— conjuntos de datos lo suficientemente completos como para analizarlos o ensayar distintas estrategias de exploración de las superficies de respuesta. En particular, encontramos que sí existen varios casos de análisis del tipo “estático”, en los cuales se explica cómo se calculan los efectos de los factores de variabilidad para modelos de regresión con *DB*, que hemos presentado sintéticamente en el capítulo anterior. Asimismo, también se encuentran disponibles algunos trabajos sobre aplicaciones de estos modelos a los casos de diseños factoriales a dos niveles, en particular, mediante modelos de primer orden, también comentados en el capítulo anterior.

Por ello, encontramos en el desarrollo de este trabajo al menos dos dificultades concretas: (a) salvo unas muy pocas publicaciones y el conocido enfoque “clásico”, prácticamente no hay antecedentes sobre este tipo de estudios para *MDB*, y (b) la imposibilidad práctica de obtener datos de casos reales para analizar.

Frente a estas restricciones relativamente importantes, intentaremos poner un poco de claridad en ambos puntos a través de las siguientes propuestas de trabajo:

- a. De cara a la falta de antecedentes, consideramos adecuado presentar en este capítulo un enfoque conceptual del problema de estudio que, aunque introductorio, servirá para ubicar mejor el problema, al menos desde un punto de vista de la visualización del mismo. De este modo, con esta presentación del tema, creemos puede resultar útil una introducción visual al tema, que tome como punto de

¹ *Vid.*, p. ej., Box *et al.* (2005), p. 526.

partida a la teoría existente (diseños factoriales para *DB*, o enfoque estático), y le incorpore el resto de los elementos clave que componen la *MSR*: criterios de paso de un diseño de primer orden a otro (p. ej.: mediante la estrategia de la máxima pendiente de crecimiento), definición de criterios de ajuste de modelos de primero y segundo orden y definición de diseños para modelos de segundo orden, entre otros. En este capítulo haremos esta introducción a partir de una descripción cualitativa de los diseños estáticos y de cómo se va construyendo el enfoque secuencial o enfoque dinámico.

- b. Debido a que el experimento que proponemos no se ha realizado hasta el momento —y valga la aclaración— recurriremos a modelos simulados aleatoriamente mediante la ayuda del paquete estadístico **R**, que sí nos permitirá disponer de datos en cada una de las etapas del enfoque secuencial, simulando el experimento. Los detalles de esta parte, los abordaremos en el capítulo siguiente, como así los de sus fundamentos principales.

4.3. Breves notas históricas

Sin extendernos demasiado en estos aspectos, creemos que resultará útil puntualizar los hitos más importantes que ha atravesado el desarrollo de la *MSR*, desde sus orígenes hasta la actualidad, como contexto gradual de avance del conocimiento de este tema. Esto nos permitirá definir con mejor claridad qué oportunidades de estudio definen las preguntas que aún no se han contestado con la misma profundidad que en el caso de los modelos clásicos, lo que nos dará lugar a justificar de alguna forma la intención que perseguimos con nuestro presente trabajo, siempre con el enlace *MSR-MLG* como perspectiva principal.

4.3.1. Modelos clásicos de *MSR*

Desde los trabajos iniciales acerca de la *MSR* durante los 1950 hasta casi fines de la década de 1980, el desarrollo de este tema alcanzó un grado muy avanzado y autorizado de madurez, tanto en los aspectos y desarrollos metodológicos como en las aplicaciones prácticas, tanto en problemas técnicos como de otras ramas de la ciencia. En trabajos recopilatorios como los de HILL Y HUNTER (1966); MYERS *et al.* (1989) y MYERS *et al.* (2004) puede encontrarse un detalle pormenorizado de la extensa cantidad de autores que han realizado contribuciones en este sentido. La característica subyacente de estos primeros desarrollos ha rondado principalmente alrededor de modelos clásicos, cuya parte sistemática es de naturaleza lineal y cuya parte aleatoria sigue una distribución normal. Los factores de variabilidad han sido considerados fijos

y el tipo de proceso bajo estudio se ha estudiado tanto para problemas con una sola respuesta como así también para casos multirrespuesta. Para el caso de modelos del tipo no lineal pero transformables en lineales, se han estudiado también sus aspectos y consecuencias. El artículo de MEAD Y PIKE (1975) hace también una revisión histórico-metodológica muy completa del desarrollo de la *MSR* y sus métodos relacionados, con numerosas referencias y antecedentes, haciendo una curiosa división de aportaciones las cuales nombran como “Before Box” y “The Box Revolution”, que hace obvia referencia al trabajo de BOX Y WILSON (1951).

En los trabajos de MYERS *et al.* (1989 y 2004) los autores discuten problemas tales como la optimalidad de los diseños y los diseños para mezclas, entre otros. Tomando como punto de partida el trabajo de BOX Y LUCAS (1959), también abordan los autores el problema de los modelos no lineales para *MSR*, destacando el problema típico de esta clase de modelos, que es el de conocer valores iniciales para el vector de parámetros del modelo². Dicho de otro modo, estos trabajos muestran la natural dificultad de llegar a diseños óptimos para modelos no lineales debido a la dependencia de la matriz de varianzas y covarianzas en el vector de parámetros. Si bien el enfoque para resolver este problema no es “exacto” —cuanto menos, en el sentido que tiene el mismo en el modelo lineal normal— sino “iterativo”, existen trabajos dedicados en profundidad al mismo, varios de los cuales se citan en el mencionado artículo. Este mismo problema es también discutido en MEAD Y PIKE (1975).

4.3.2. Enfoque estático y *MLG*

Resulta claro que la distribución que siguen los datos bajo estudio tiene un impacto importante en la estimación de los parámetros del modelo, como así también en las inferencias que se realicen a partir de ellos y, particularmente, a partir del problema del *análisis* de las superficies de respuesta, que es el problema natural que sigue al del *diseño*. En MYERS *et al.* (1989), por ejemplo, se citan trabajos en donde se han estudiado modelos de la familia exponencial aplicados a la *MSR*, tales como los de PAYNE³ y los de SOLANA y colaboradores⁴, entre otros. En MEAD Y PIKE (1975) también se describen trabajos que refieren a las primeras incursiones en estudios de *MSR* que siguen distribuciones distintas a la normal, como es el caso del uso del

²Resulta muy contundente e ilustrativa en este sentido la frase atribuida a William G. Cochran (1909-1980): “*You tell me the value of θ and I promise to design the best experiment for estimating θ* ”. [Fuente: COCHRAN, WILLIAM G. (1973). “Experiments for Nonlinear Functions”. *Journal of the American Statistical Association*, N° 68, pp. 771 *et seq.* Citado en MYERS *et al.* (1989)].

³PAYNE, C. D. (1985). *The GLIM Systems Release 3.77*. Oxford. Numerical Algorithms Group.

⁴SOLANA, R. P.; CHINCHILLI, V. M.; WILSON, J. D.; CARTER, W. H. y CARCHMAN, R. A. (1986). “Estimation and Analysis of the Concentration Response Surfaces Associated With Multiple Agent Combinations”. *Toxicology and Applied Pharmacology*, N° 85, pp. 231 *et seq.*

modelo *probit* para modelar proporciones o el de varias funciones de aproximación de naturaleza polinómica (*op. cit.*, pp. 817 *et seq.*). Sin embargo, la mayor parte de estos antecedentes se concentra casi en su totalidad en el problema del *análisis*, no desarrollando en profundidad el problema del *diseño*.

Es tal vez en el trabajo de MYERS *et al.* (1989) en donde se comienza a mencionar el vacío metodológico que existe en intentar realizar diseños de experimentos secuencialmente —en el sentido de la *MSR*— para problemas en los cuales la respuesta sigue alguna de las distribuciones de la familia exponencial que sea distinta a la normal. Y es —quizá— en esta misma referencia en donde por primera vez se puntualiza la posibilidad de enlazar este problema a partir de las posibilidades de la formulación dada en NELDER Y WEDDERBURN (1972) sobre *MLG*. También se señalan en la misma referencia algunas posibles direcciones de futuras investigaciones que resultan interesantes.

Es importante destacar que estas contribuciones realizadas en procura de “enlazar” los *MLG* con los diseños factoriales —y de acuerdo con la terminología que venimos empleando— haría referencia al *punto de vista estático* del problema de la *MSR* no clásica, es decir, con respuestas que siguen distribuciones distintas a la normal. Es preciso destacar, entre otras valiosas contribuciones, las de PUJOL (1990), MYERS *et al.* (1994), HAMADA Y NELDER (1997), LEWIS (1998), LEWIS *et al.* (1999, 2001A y 2001B), WU Y HAMADA (2000) y LEWIS *et al.* (2001A y 2001B).

4.3.3. Enfoque dinámico, *MSR* y *MLG*: una revisión crítica

Muy posiblemente sean los trabajos de KHURI (1993 Y 2001) en donde se especifique puntualmente lo que venimos dando en llamar el *punto de vista dinámico* del problema de la *MSR*, casi con el mismo espíritu con que nos hemos propuesto estudiarlo en la presente tesis. El problema de la naturaleza no lineal de superficies que sigan distribuciones de la familia exponencial no normales es presentado como una de las naturales dificultades para la estimación de los parámetros del modelo, como así también para la planificación de diseños óptimos. El autor hace una breve revisión histórica de las principales aportaciones en este sentido, tanto antes como después del trabajo de NELDER Y WEDDERBURN (1972). Asimismo, el autor enfatiza que el marco metodológico propuesto por los *MLG* indica que la clasificación de factores tanto cualitativos como cuantitativos en los problemas de *MSR* no constituye un problema, como así tampoco la no normalidad y la falta de constancia en la varianza (*op. cit.*, p. 194 y 195). En el mismo artículo se esboza el enfoque que podría seguirse para aumentar la probabilidad de supervivencia de pacientes que sufren cierto tipo de cáncer. El tema de la optimalidad de los diseños para esta clase de problemas también es abordado: el

autor parte del enfoque clásico de los diseños alfabéticamente óptimos, indicando con claridad la dificultad de encontrar una matriz de “pesos” \mathbf{W} que minimice la matriz asintótica de varianzas y covarianzas de los parámetros debido a la dependencia de \mathbf{W} en éstos y en la forma de definir la función link.

En nuestra opinión, diremos que el problema del enlace *MSR-MLG* se encuentra adecuadamente formulado y claramente enunciado en estas publicaciones del Prof. Khuri, destacando al final de la misma la necesidad de continuar estudiando el problema de la naturaleza no lineal de esta familia de modelos, especialmente cuando se pretenda abordar el tema de la optimalidad de los diseños (*op. cit.*, p. 201).

Siguiendo una línea de trabajo coherente, en los capítulos 10 y 12 de KHURI Y CORNELL (1996) se amplía el enfoque, aportaciones y alternativas disponibles para el estudio de modelos no lineales para *MSR*, tanto desde el enfoque estático como así también del dinámico, aunque menos desarrollado que el anterior, y volviendo a puntualizar la necesidad de continuar investigando el enlace *MSR-MLG*.

En MYERS (1999) y en MYERS *et al.* (2004) se analiza con elegancia el estado del arte y futuras direcciones de investigación sobre *MSR*, con numerosos y valiosos comentarios y discusiones de calificados estudiosos de la materia. Los autores dedican varias páginas a discutir problemas importantes tales como la optimalidad de los diseños, detalles del enfoque bayesiano del problema, como así también una ilustración sobre ciertas implicaciones en la consideración de diseños óptimos para el modelo logístico univariante y para el de Poisson. Los comentarios y opiniones del Prof. Khuri al artículo precitado (*op. cit.*, pp. 58 *et seq.*) son especialmente valiosas en cuanto al sentido de estudio que venimos siguiendo en esta tesis, sin que por ello el resto de las contribuciones posean una valoración menor.

4.3.4. Una perspectiva sobre el estado actual del arte

Las publicaciones más recientes que han aparecido al respecto del enfoque dinámico precitado son KHURI (2001); MYERS Y MONTGOMERY (2002) y MYERS *et al.* (2004). En todas ellas, se hace especial hincapié en la obtención de diseños óptimos para *MLG*. En la primera de ellas, se dedica una sección particular a la elección de diseños óptimos, en donde se introducen brevemente cuatro enfoques para lidiar con el problema de la dependencia de la matriz asintótica de varianzas y covarianzas de los parámetros en éstos, asociado con el estudio de la optimalidad:

- a. *El enfoque secuencial*, desarrollado inicialmente en WU (1985),
- b. *El enfoque bayesiano*, desarrollado en CHALONER Y LARNTZ (1989)⁵, y

⁵CHALONER, K. y LARNTZ, K. (1989). “Optimal Bayesian Design Applied to Logistic Regression Experiments”. *Journal of Statistical Planning and Inference*, N° 21, pp. 191 *et seq.*

- c. *El enfoque minimax*, desarrollado en SITTER (1992).
- d. *El enfoque “gráfico comparativo”*, desarrollado en ROBINSON Y KHURI (2000)⁶.

El mismo artículo contiene una reflexión que hemos considerado apropiada citar textualmente para resumir el *statu quo* de estos temas: “*The majority of designs for GLMs are for logistic models with only one control variable and two parameters in each model. Practitioners would feel that the study is not truly an RSM one unless there are at least two control variables*” (*op. cit.*, p. 2027), que podría ser traducido sin mucho error como: “La mayoría de diseños para *MLG* son para modelos logísticos, con solamente una variable de control y dos parámetros en cada modelo. Los estudiosos podrían sentir que este estudio no es verdaderamente un problema de *MSR* a menos que haya dos variables de control”. En este sentido, en MYERS (1999) pueden encontrarse casos interesantes, como la discusión sobre la optimalidad de diseños para modelos logísticos con 2 variables de control y una visión útil sobre respuestas que siguen una distribución de Poisson. El artículo termina proponiendo precisiones y formulaciones para el problema de los diseños óptimos para *MLG*.

En el más recientemente publicado libro sobre *MSR*, MYERS Y MONTGOMERY (2002), los autores dedican un capítulo entero (el número 10) al enfoque estático en general, introduciendo los *MLG*, los problemas de la elección de las funciones “link”, los problemas en la estimación de parámetros, etc. Complementariamente, en la obra más reciente referida a *MLG*, de MYERS *et al.* (2002), se continúa dando un enfoque marcadamente más estático que dinámico, del cual prácticamente no se introducen ni discuten nuevas aplicaciones más que las ya mencionadas en párrafos anteriores de nuestro trabajo.

En resumidas cuentas, y de acuerdo con la revisión bibliográfica tanto clásica como actual que hemos revisado y escudriñado con mucha curiosidad, podemos finalizar estas breves notas históricas indicando cuatro conclusiones muy sencillas, pero que pueden sintetizar un punto de vista sobre el *estado del arte* de nuestro tema de investigación al momento de prepararlo:

1. El problema de construir diseños factoriales para ajustar modelos estadísticos que sigan distribuciones normales, goza de un estado de desarrollo razonablemente maduro, habiendo importantes publicaciones en las que pueden encontrarse tanto aportaciones metodológicas como estudios de casos reales resueltos con éxito. Ello no obstante, y como hemos comentado anteriormente, no queda actualmente

⁶ROBINSON, K. y KHURI, ANDRE I. (2000). *Quantile Dispersion Graphs for the Comparison of Designs for Logistic Regression Models*. Technical Report. Department of Statistics, University of Florida, Gainesville.

clara la conveniencia en el uso de diseños factoriales por encima de otros para el caso de distribuciones de la familia exponencial, si bien se ha utilizado en algunas ocasiones⁷. Diremos así, que el *punto de vista estático* del problema de la *MSR* para *MLG* tiene una “inercia” muy favorable, la cual motivará seguramente a los investigadores a continuar con su refinamiento metodológico y con su expansión hacia nuevas aplicaciones.

2. El encadenamiento de diseños estáticos en el sentido de la experimentación secuencial, es decir, el *punto de vista dinámico*, parece estar en una etapa de madurez de inferior desarrollo que el estático en lo que refiere a *MLG*. Los trabajos publicados hasta la actualidad son relativamente escasos, y creemos que es conveniente continuar esta chispa inicial de los autores que lo han abordado, de forma que pueda conseguirse una llama de luz más consistente, que permita aportar una mayor claridad a este problema. Aquí es a donde mayormente apunta nuestro trabajo. Si bien el problema de la *estimación* de parámetros y el de *análisis* de superficies de respuesta pertenece al enfoque dinámico, el desarrollo de estos aspectos se encuentra en mejores condiciones metodológicas que el problema del *diseño*, es decir, de la forma de cómo elegir los puntos de diseño en cada una de los diseños estáticos que configuran el enfoque dinámico. Sobre este punto preciso, y mediante el desarrollo del presente trabajo, es que hemos realizado el esfuerzo por intentar ver con un poco más de claridad este aspecto, de forma de plantear nuevas perspectivas para observar y estudiar el mismo y que sirvan para continuar agregando valor al nivel de conocimiento sobre estos temas.
3. Uno de los puntos en donde más se ha avanzado, con líneas concretas ya trazadas, es el del estudio de los *diseños óptimos*. Este es otro de los puntos que goza de buen estado de desarrollo, si bien las aportaciones realizadas en este tema puntual son relativamente recientes.

4.4. Ejemplo: diseños con dos factores

El problema general, aquí podría contextualizarse dentro de una superficie de respuesta $\pi(x_1, x_2, \beta)$, de parámetros desconocidos, que representa un cierto proceso de la realidad. De dicha superficie se desea mediante la experimentación secuencial poder encontrar mejores condiciones operativas, ya sea que conduzcan hacia un máximo, hacia un mínimo o bien hacia un valor objetivo. Para el caso de una cierta región de operabilidad \aleph definida por el rectángulo de vértices $abcd$, podemos representarla conceptualmente en la figura 4.1:

⁷ Vid. p. ej. MYERS *et al.* (2002).

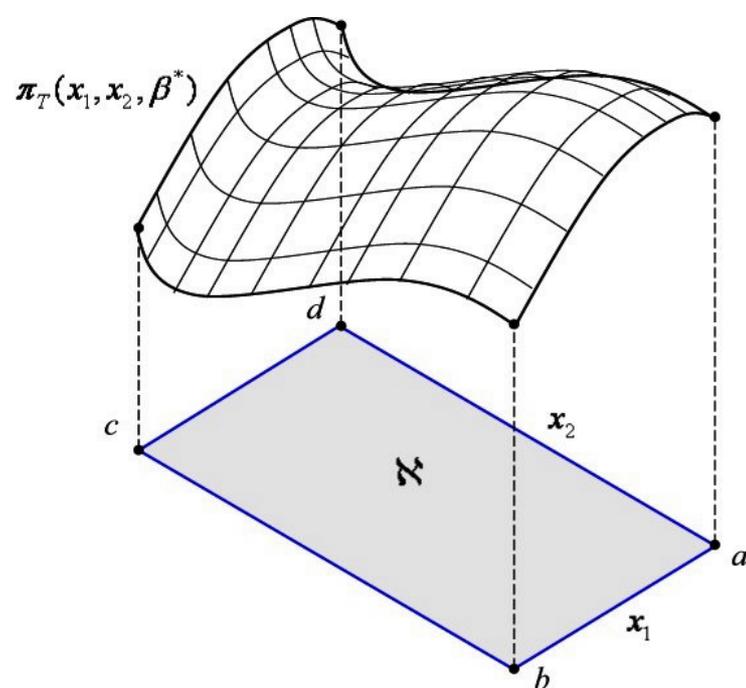


Figura 4.1: Una idea de proceso real que se presupone depende de dos factores x_1 y x_2 , representado mediante una superficie de respuesta $\pi_T(x_1, x_2, \beta^*)$, en la que el subíndice T refiere a que es de carácter “teórica” en un principio, y el superíndice del vector β refiere a su carácter desconocido.

Al considerar al valor esperado de la respuesta como una función no lineal a dos factores, se tendrá el mismo problema que se tenía con un solo factor (ver capítulo 3). La dependencia de esta función en los parámetros, exigirá que se realicen estimaciones iterativas de estos últimos. En la figura 4.2 se ilustra este caso con un plano de aspecto sigmoide, y en la figura que le sigue (4.3) puede apreciarse el aspecto que tienen las curvas de nivel de aquélla sobre el plano de los factores.

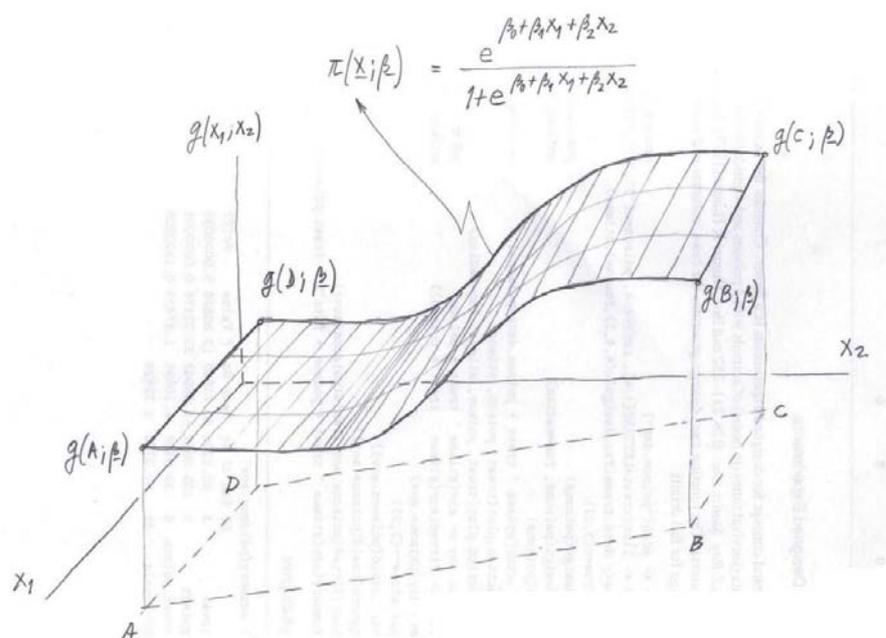


Figura 4.2: Representación de una respuesta en función de dos factores de variabilidad, que sigue una distribución logística. Sobre alguna parte de la región $ABCD$ habrá que definir los puntos de diseño que permitan conocer $\pi(\mathbf{x}, \beta)$

La consideración de la proyección de $\pi(x_1, x_2, \beta)$ da una idea bastante clara que la “densidad de la información” contenida en la proyección sobre el plano de los factores no será tan clara antes de realizar el experimento. Puede verse en la 4.3, por ejemplo, que en las zonas delimitadas por los rectángulos $D'A'A''D''$ o bien por el simétrico $C'B'B''C''$, la “densidad de información” que ofrece el modelo teórico es diferente a la que aparece en los rectángulos extremos $DAA'D'$ y $CBB'C'$, a partir de lo cual nuevamente queda claro que no será posible la consideración de diseños óptimos para estos casos. El buen sentido indicaría que, si se supone que el modelo considerado tiene naturaleza similar a la de este esquema, habrá que invertir un mayor esfuerzo experimental dentro de las zonas con mayor densidad de información que en las extremas. Como esto no se puede saber de antemano, la consideración de diseños secuenciales

resulta inevitable.

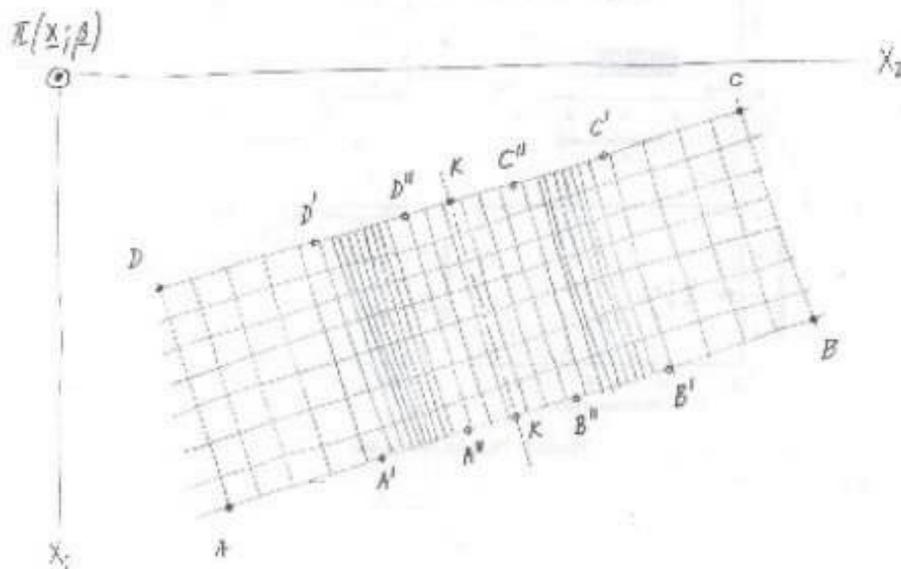


Figura 4.3: Proyección de la superficie no lineal a dos niveles sobre el plano de los factores, formando una figura rectangular de vértices $ABCD$.

4.5. El enfoque dinámico de diseños

4.5.1. Introducción

En las páginas anteriores hemos tratado de introducir de un modo conceptual qué papel juegan los diseños, viendo éstos como “fotografías” o enfoques de carácter “estático” de la evolución del avance del conocimiento sobre el sistema desconocido. En esta sección haremos un primer esbozo comentado sobre diseños particulares a utilizar, que serán los de primero y segundo orden.

Primero, mostraremos también de modo cualitativo de qué manera se van encadenando los diseños de primer orden cuando se supone que la región de experimentación inicial se encuentra lo suficientemente alejada del óptimo local, cuanto menos, del que se podría alcanzar dentro de las capacidades técnicas y económicas que acoten las condiciones de experimentación. Indicaremos también de forma cualitativa, cómo se desplazarán estas regiones en busca de zonas de mayor valor de la respuesta —en el caso de maximización— para lo cual utilizaremos una herramienta del enfoque clásico: el criterio de la “steepest ascent”.

Y en segundo lugar, cuando se den las condiciones de falta de ajuste del modelo

de primer orden, se plantearán modelos de segundo orden que tengan en cuenta la curvatura que supuestamente estará presente en la región de experimentación.

A todo esto trataremos de introducirlo de modo visual-conceptual, y dejamos su desarrollo teórico para el capítulo siguiente, en donde nos adentraremos de lleno en los detalles más salientes de este nuevo enfoque que presentamos.

4.5.2. Funciones de aproximación

Un supuesto bastante razonable del que suele partirse cuando se comienza a experimentar es que se tiene alguna cierta evidencia de que las condiciones actuales de operación de un proceso no son las óptimas y que se supone que será posible encontrar algunas condiciones mejores que las ordinarias, tales que hagan que puedan lograrse “mejores” valores de la variable que se pretende estudiar. En términos geométricos, esto puede pensarse como una región de la superficie de respuesta de los factores, $f(\boldsymbol{\xi}, \boldsymbol{\theta})$, que no presente rugosidades⁸ en las inmediaciones de la región de operación actual, con lo cual, diremos que sería de esperar que al comienzo de la experimentación, la superficie de respuesta presente aspectos “planos”. Esto nos llevará a pensar que una primera aproximación a dicha superficie, localmente, pudiera ser un modelo de primer orden.

Partimos entonces de la construcción de funciones de aproximación de primer orden: rectas, planos e hiperplanos, de tal manera que ellas resulten —mientras no se tenga evidencia de lo contrario— suficientes para conocer la forma de la superficie en el comienzo de las pruebas. En la figura siguiente (figura 4.4) ilustramos esta situación:

Con los puntos así generados —todos con subíndice G —, se estimarán los parámetros del modelo de aproximación ($p = 3$), con los cuales se determinará el modelo de aproximación ajustado, que será de la forma:

$$\hat{\pi}(\mathbf{x}, \boldsymbol{\beta}) = \frac{\exp(\hat{\beta}_0 + \hat{\beta}_1 x_1 + \hat{\beta}_2 x_2)}{1 + \exp(\hat{\beta}_0 + \hat{\beta}_1 x_1 + \hat{\beta}_2 x_2)}$$

De este modo, se consigue un modelo de aproximación para la superficie verdadera y desconocida, $f(\mathbf{x}, \boldsymbol{\beta})$, dentro de la región inicial de experimentación que, al suponerse válida la hipótesis según la cual esta región se encuentra lo suficientemente lejos del óptimo, esta aproximación puede resultar razonable.

4.5.3. Secuencia de aproximaciones

Una representación cualitativa de una superficie teórica desconocida, $f(\mathbf{x}, \boldsymbol{\beta})$, puede verse en la figura 4.5, que nos servirá de guía para la aproximación analítica posterior

⁸ Vid., p. ej., PRAT *et al.* (2004), p. 165.

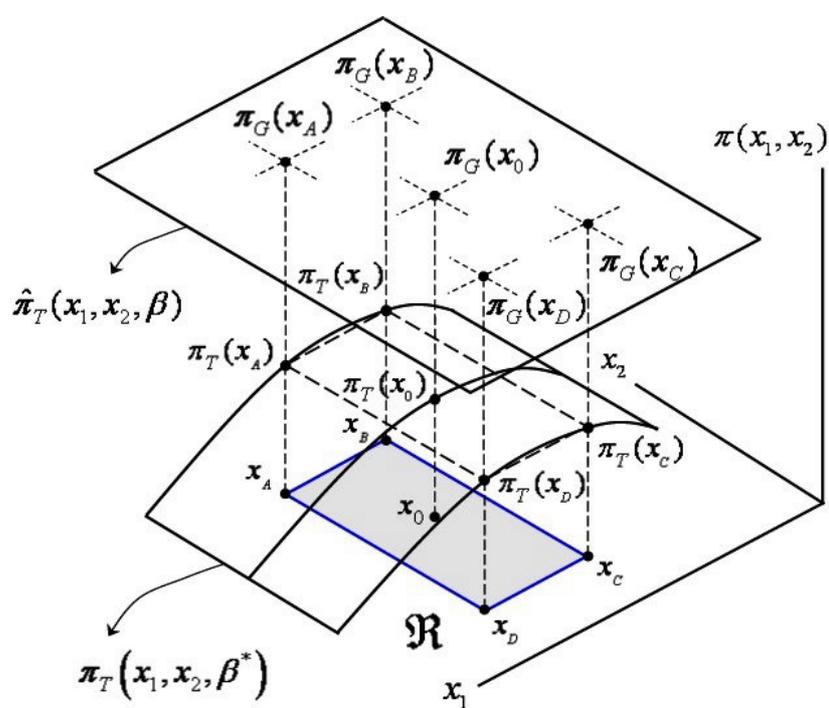


Figura 4.4: Idea de aproximación de la superficie teórica mediante un modelo ajustado de primer orden.

que ensayaremos en el capítulo siguiente. Como venimos comentando desde secciones anteriores, asociaremos a dicha superficie una probabilidad de éxito, la que estará acotada superiormente por el plano $\pi(\mathbf{x}, \boldsymbol{\beta}) = 1$, e inferiormente por $\pi(\mathbf{x}, \boldsymbol{\beta}) = 0$, ambos paralelos al plano de los factores.

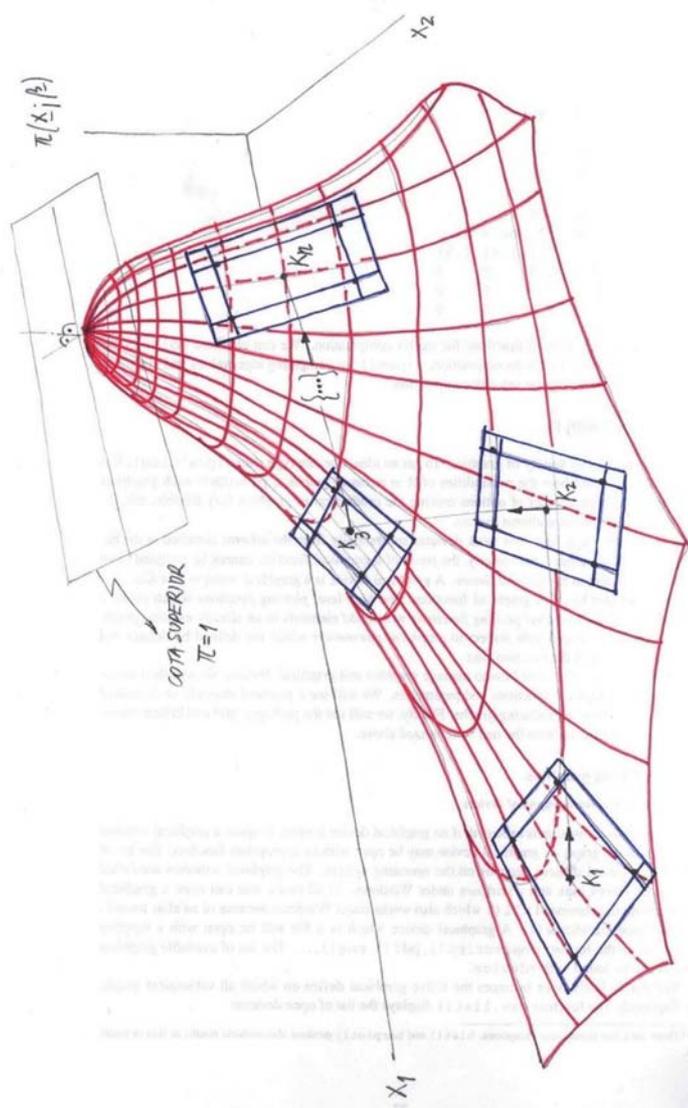


Figura 4.5: Esquema de aproximaciones secuenciales mediante modelos de primer orden.

Como indicáramos anteriormente, cuando nos encontramos “inicialmente lejos” de las condiciones óptimas, se ensayarán aproximaciones de primer orden. En un primer diseño (ver figura 4.5), cuyo centro diremos que es K_1 , el procedimiento clásico de *MSR* pretende buscar un segundo diseño, de centro K_2 , que se pueda definir a partir

del primero. Para ello, el método clásico propone encontrar un vector que resulte colineal con el gradiente de la superficie ajustada, $\pi(\mathbf{x}, \hat{\boldsymbol{\beta}})$, y que defina una dirección de máximo crecimiento en los factores.

De acuerdo con lo anterior, el problema se concentra en hallar las soluciones en $\boldsymbol{\beta}$ de las ecuaciones:

$$\begin{aligned} \blacksquare \frac{\partial}{\partial x_1} [\pi(\mathbf{x}, \hat{\boldsymbol{\beta}})] &= 0, \text{ y} \\ \blacksquare \frac{\partial}{\partial x_2} [\pi(\mathbf{x}, \hat{\boldsymbol{\beta}})] &= 0, \end{aligned}$$

de modo que pueda calcularse un cierto vector $\mathbf{s} = (s_1, s_2)'$ que indique la dirección de máximo crecimiento de la superficie, sobre la que se definirá el diseño de centro K_2 .

Siguiendo esta metodología, se continuarán las aproximaciones sucesivas⁹ hasta llegar a un cierto diseño, de centro K_n , a partir del cual el modelo estimado de primer orden no evidencie un buen ajuste, lo cual será indicativo de que será necesario incluir ya sea términos cruzados del tipo x_1x_2 o bien términos cuadráticos en el modelo propuesto. Cuando tengamos evidencia significativa de este hecho, los modelos de primer orden propuestos anteriormente tendrán que ser complementados con los términos que mencionamos, con los cuales se pasará a considerar *modelos de aproximación de segundo orden* para $f(\mathbf{x}, \boldsymbol{\beta})$.

En la figura siguiente (figura 4.6) puede observarse una ilustración que representa el proceso de aproximación de primer orden que ocurre en el plano de los factores, en donde las curvas cerradas del tipo $\pi = \text{cte.}$ serán las curvas de nivel de $f(\mathbf{x}, \boldsymbol{\beta})$.

El experimentador se encuentra ahora con que —seguramente debido a la presencia de curvatura en la zona de experimentación— ya no le será de mucha utilidad planear nuevos diseños para modelos de primer orden, sino que más bien pensará en considerar diseños que le permitan ajustar funciones de aproximación cuyo predictor lineal más general sea del tipo:

$$E[\pi(\mathbf{x}, \boldsymbol{\beta})] = \beta_0 + \beta_1x_1 + \beta_2x_2 + \beta_{12}x_1x_2 + \beta_{11}x_1^2 + \beta_{22}x_2^2$$

Un esquema de esta situación puede verse en la figura 4.7: el último diseño de primer orden que evidencia un buen ajuste es el que se representa con centro K_n y de ecuación genérica $g_1(\mathbf{x}, \hat{\boldsymbol{\beta}})$, y por la naturaleza de falta de ajuste, el diseño siguiente —que se supondrá que corresponde a zonas con curvatura dentro de la superficie— deberá contemplar distintos lay-outs a los anteriores. Un ejemplo de esto se esquematiza definiendo un típico diseño “estrella”, de vértices $A_{n+1}B_{n+1}C_{n+1}D_{n+1}$, de centro K_{n+1}^0 , y cuyo “mástil de estrella” (ver detalle en la figura) estará formado por los puntos K_{n+1}^S, K_{n+1}^0 y K_{n+1}^I :

⁹Siempre y cuando las condiciones técnicas y presupuestarias lo permitan.

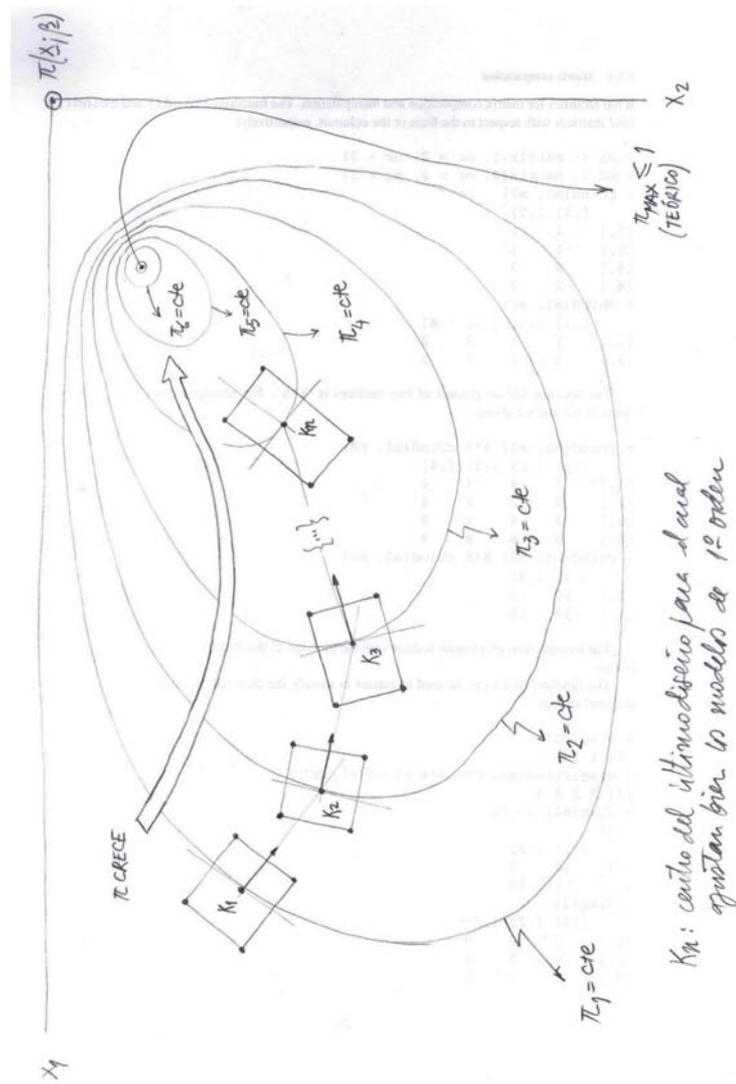


Figura 4.6: Situación en el plano de los factores de las aproximaciones sucesivas con modelos secuenciales de primer orden.

A partir de este diseño de segundo orden, se construirá una nueva función de aproximación $g_2(\mathbf{x}, \widehat{\boldsymbol{\beta}})$ que contemple la curvatura de la superficie teórica, $f(\mathbf{x}, \boldsymbol{\beta})$, de forma de aproximarse con mayor eficiencia al potencial máximo local y describiendo la superficie de la manera más precisa posible. En nuestro desarrollo propuesto en este trabajo, utilizaremos toda la información acumulada. Para ello, partimos del el primer diseño definido en la zona de aproximación lineal y llegamos al último de todos, en la zona que evidencia curvatura, luego de lo cual, el presupuesto disponible ya se haya acabado. Aprovechando toda esta información, es que ajustaremos modelos sucesivos que vayan utilizando toda la información acumulada disponible en los mismos.

4.6. Superficies de respuesta para *DB*

4.6.1. Generalidades

Siguiendo el esquema de razonamiento comentado para datos con distribución normal, vemos aquí qué aspecto tienen las superficies de respuestas para datos provenientes de distribuciones binarias, a partir de lo cual, dejaremos trazada la base desde la cual estudiaremos las superficies de respuesta que de ellos provengan.

En términos generales, diremos que nuestro foco estará puesto en una cierta superficie “verdadera” o “teórica” dada por un cierto predictor lineal $g[\pi(\mathbf{x}, \boldsymbol{\beta})]$, que como vimos, será de la forma:

$$g[\pi(\mathbf{x}, \boldsymbol{\beta})] = \beta_0 + \mathbf{x}'\mathbf{b} + \mathbf{x}'\mathbf{B}\mathbf{x}.$$

En el caso más general, esta función será un polinomio completo de orden k , con todos los términos de interacción que correspondan.

De acuerdo con la estrategia secuencial de experimentación, se pretende definir funciones de aproximación en los diseños sucesivos que permitan “tomar fotografías” de la superficie de respuesta en sus diferentes etapas de exploración, para configurar un cierto “enfoque dinámico” con el cual estudiar dichas superficies.

4.6.2. Algunos ejemplos

A modo ilustrativo, mostraremos en esta última sección algunos ejemplos de representaciones interesantes de superficies obtenidas al darle otros valores al vector de coeficientes $\boldsymbol{\beta}$, como así también las respectivas curvas de nivel en cada caso¹⁰, siguiendo la definición para $g[\pi(\mathbf{x}, \boldsymbol{\beta})]$ dada en la última ecuación. Todos estos gráficos se encuentran representados en escala logística.

¹⁰Estos gráficos los hemos realizado utilizando el software **S-Plus**, que fue con el que inicialmente comenzamos nuestro estudio.

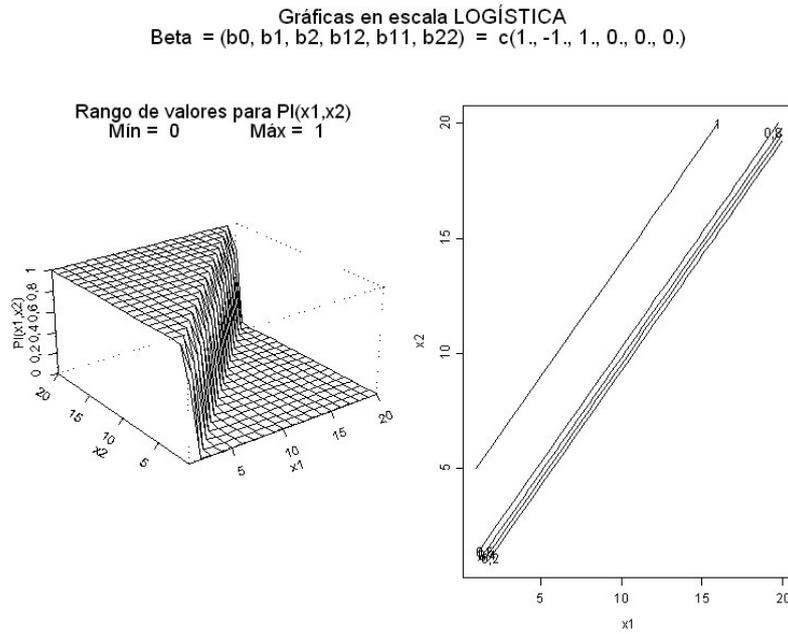


Figura 4.8: Ejemplo 1: “cascada diagonal simple” de superficie de respuesta en escala logística.

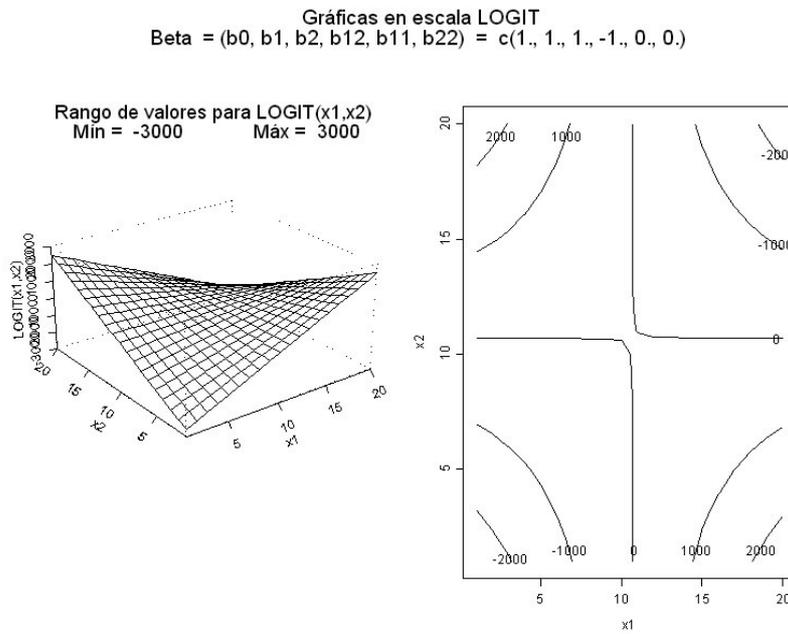


Figura 4.9: Ejemplo 2: “punto de silla” de superficie de respuesta en escala logística.

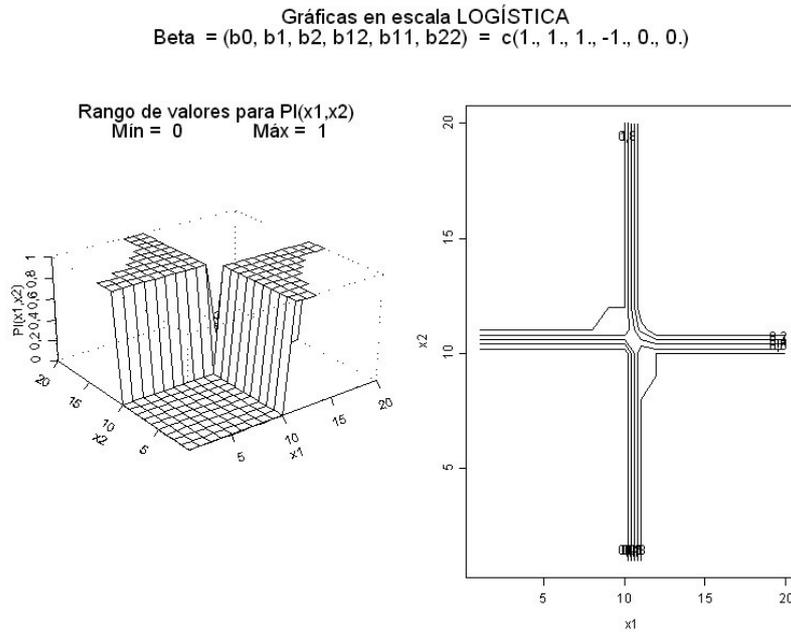


Figura 4.10: Ejemplo 3: “cascada de doble cuadrante rectangular” de superficie de respuesta en escala logística.

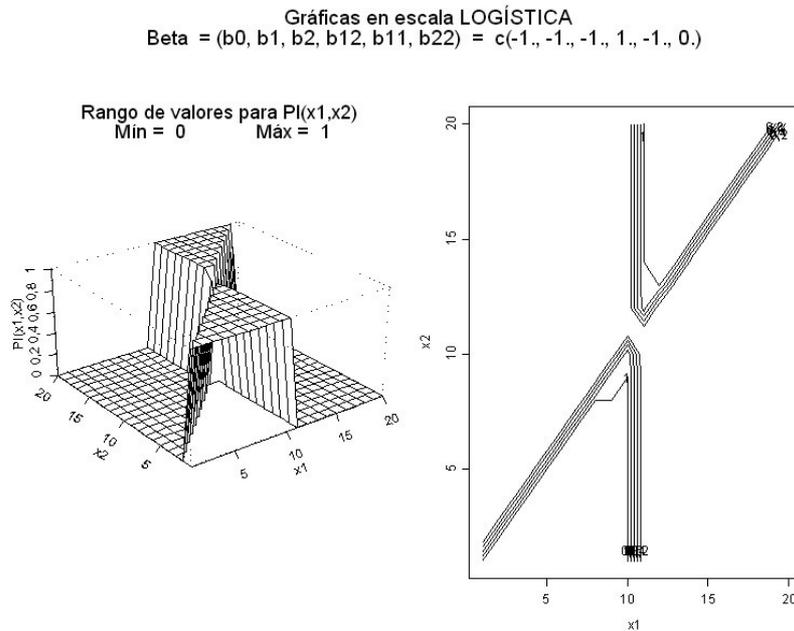


Figura 4.11: Ejemplo 4: “cascada en cruz-diagonal” o “cascada de doble cuadrante agudo” de superficie de respuesta en escala logística.

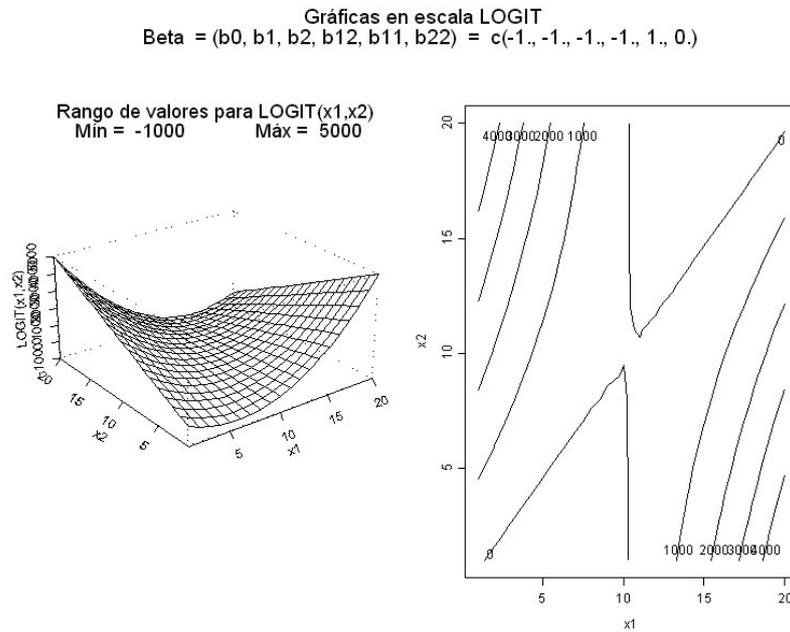


Figura 4.12: Ejemplo 5: “semi teja-cruz oblicua” de superficie de respuesta en escala logística.

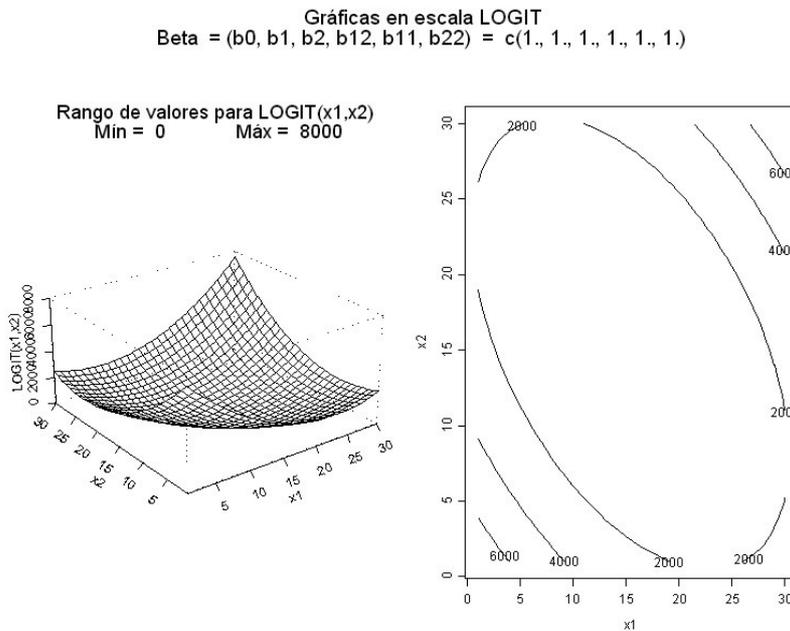


Figura 4.13: Ejemplo 6: “mínimo” de superficie de respuesta en escala logística.

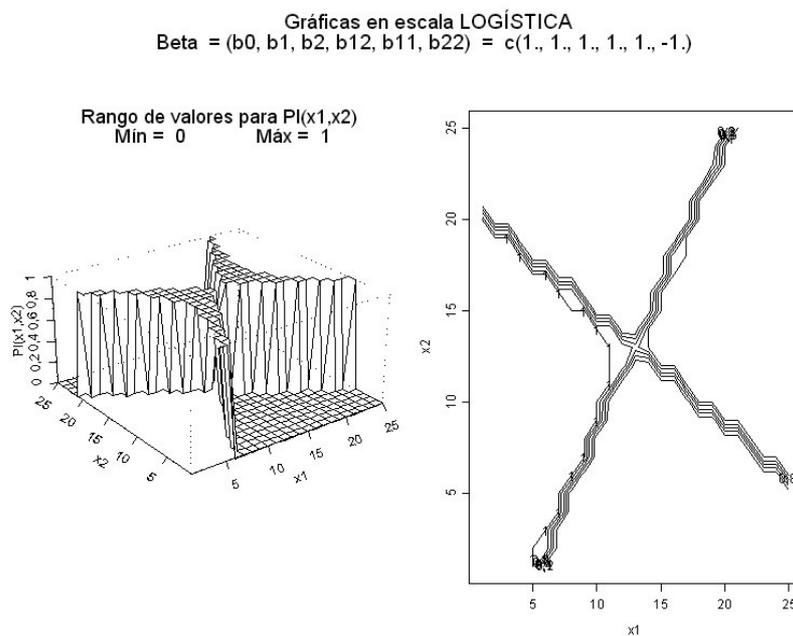


Figura 4.14: Ejemplo 7: “cascada en cruz-cruz” o “cascada doble cruz” de superficie de respuesta en escala logística.

En el capítulo siguiente, describiremos y haremos una discusión de distintas propuestas de estrategias para encadenar diseños estáticos de exploración secuencial de superficies de respuesta para DB , conformando una propuesta concreta de enfoque dinámico del problema que nos ocupa.