

Funciones test para optimización mono-objetivo

Apellidos, nombre	Carpitella, Silvia ¹ (silvia.carpitella@unipa.it) Izquierdo Sebastián, Joaquín ² (jizquier@upv.es)
Departamento	¹ Dipartimento dell'Innovazione Industriale e Digitale ² Departamento de Matemática Aplicada
Centro	¹ Università degli Studi di Palermo ² Universitat Politècnica de València

1 Resumen

Optimización es una palabra (¡y una práctica!) que permea prácticamente cualquier actividad humana. En todos los planes de estudio, especialmente en las Ingenierías, se incluyen conceptos y técnicas básicas de optimización. Con frecuencia, todo el bagaje sobre optimización suele reducirse a los elementos más simples relacionados con la derivación. Como resultado, el material sobre optimización se reduce a cuestiones muy elementales, que tan solo permiten resolver problemas de optimización muy simples e idealizados. Para resolver la mayor parte de problemas del mundo real son necesarias técnicas sofisticadas, tales como los algoritmos evolutivos, que no se basan en el Cálculo Infinitesimal y que, en consecuencia, nunca caben en tales planes de estudios. No obstante, el ingeniero tendrá que utilizar tales técnicas antes o después. Para poner a prueba la habilidad de tales técnicas de optimización se suelen utilizar problemas de benchmarking que exhiben algunas de las características de los problemas del mundo real. En este artículo enumeramos brevemente algunas de tales características y presentamos una colección de problemas de optimización mono-objetivo, no condicionada, en varias variables. Tales problemas pueden ser de utilidad para que los estudiantes de ingeniería practiquen y pongan a prueba sus algoritmos favoritos de optimización. Para cada uno de tales problemas presentamos su definición, la localización de su mínimo absoluto y, para los casos en que los problemas deban ser adimensionales, presentamos también el máximo absoluto en el compacto en el que habitualmente se define cada función.

2 Introducción

Los problemas de optimización del mundo real se caracterizan, en general, por diversas complejidades de dificultad variable. En general, la optimización mono-objetivo persigue obtener el óptimo global (alcanzado en uno o varios puntos) de una función en un dominio dado, la mayoría de las veces un conjunto compacto en \mathbf{R}^n . Tal vez, la mayor complejidad se deriva del hecho de que la mayoría de los problemas de optimización del mundo real son multimodales. Esto significa que muchos óptimos locales coexisten en el paisaje asociado a la función objetivo. En este tipo de problemas, muchos algoritmos de optimización fracasan [Izquierdo, J., 2016]. Además, también se pueden considerar otras complejidades. Por ejemplo, [Surjanovic y Bingham, 2017] clasifica las funciones de optimización mono-objetivo según sus características y formas específicas: multimodal (ya mencionada), en forma de cuenco, de planicie, de valle, con crestas/caídas abruptas, y otras.

Un enfoque común para investigar el rendimiento de los algoritmos de optimización en términos de precisión, robustez y fiabilidad, consiste en probarlos en un gran conjunto de funciones de prueba artificiales [Bossek, 2017]. Se ha diseñado una gran cantidad de tales funciones test de dos o más variables, que se pueden probar en contextos de optimización tanto mono como multi-objetivo, incluso en presencia de restricciones, para poner a prueba el rendimiento y comportamiento de los algoritmos de optimización.

En este documento proporcionamos una breve descripción y las fórmulas para varias de estas funciones de prueba. Todas son funciones mono-objetivo de n variables. Para cada función descrita mostramos el gráfico de superficie 3D, implementado a través del software R [R, 2011], obviamente para 2 variables. Las funciones se normalizan entre 0 y 1, mediante la siguiente expresión:

$$(f(x) - m)/(M - m),$$

siendo m y M , respectivamente, el mínimo y el máximo en el dominio de búsqueda (un compacto) considerado para cada función.

En caso de interés, puedes encontrar colecciones online de funciones test, tales como la librería GLOBAL [Hedar, A.-R], GAMS World [GAMS World], CUTE [Gold et al.], GO Test Problems [GO Test Problems]. Una exhaustiva lista de hasta 175 funciones diferentes se puede encontrar en [Jamil y Yang, 2013].

Complementamos este artículo con la sugerencia de uso de un [objeto de aprendizaje](#) que muestra algunas dificultades, con frecuencia insospechadas, que plantea la optimización. En el objeto de aprendizaje podrás utilizar varias formas de visualización de algunas de las funciones que describimos en este artículo. En cada caso, empieza con el 'enunciado', pasa después a la 'sugerencia' y, finalmente, visualiza la 'solución'. Puedes elegir también entre visualizar superficies 3D o curvas de nivel. La utilización de este objeto de aprendizaje te presentará algunas de las sorpresas que esconden los problemas de optimización del mundo real. Tal objeto de aprendizaje es de acceso libre y se encuentra en el repositorio Riunet de la Universitat Politècnica de València [Izquierdo, J., 2009].

Tras enumerar los objetivos de este artículo, la sección de desarrollo presenta los requisitos para leerlo con aprovechamiento, y pasa directamente a la presentación, por bloques, de las funciones que constituyen el núcleo del artículo. El artículo acaba con un cierre y con la lista de las referencias utilizadas.

3 Objetivos

Tras concluir con la lectura de este documento, serás capaz de:

- Identificar y explicar la complejidad intrínseca de la optimización en los problemas del mundo real relacionada con algunas de las características de la optimización.
- Poner a prueba algoritmos de optimización de que puedas disponer para evaluar su capacidad de obtener los extremos de las funciones estudiadas.

4 Desarrollo

Para leer este artículo bastará que tengas nociones de:

Requisitos
1. Funciones elementales usuales.
2. Representación gráfica de superficies en el espacio.
3. Máximos y mínimos de funciones de varias variables.
4. Conceptos algebraicos básicos.
5. Manejo básico de alguna técnica de optimización.

Tabla 1. Requisitos básicos

FUNCIONES MULTIMODALES

Como se explica en [Izquierdo, J.], son funciones que poseen numerosos extremos locales en los que frecuentemente muchos algoritmos de optimización quedan atrapados. En esta sección te presentamos cuatro de las más famosas funciones del benchmarking que ponen en aprietos, especialmente, a los algoritmos basados en la técnica del gradiente.

La función de Ackley

La función de Ackley se define habitualmente en el hipercubo $[-a, a]^n$, y su expresión es:

$$f(\mathbf{x}) = -20 \exp\left(-0.2 \sqrt{\frac{1}{n} \sum_{i=1}^n x_i^2}\right) - \exp\left(\frac{1}{n} \sum_{i=1}^n \cos(2\pi x_i)\right) + 20 + \exp(1).$$

Vista de lejos, presenta una amplia zona aparentemente plana con un agujero en el centro, donde tiene su mínimo global $m = 0$ en $\mathbf{x} = (0, \dots, 0)$. Encuentra esta función en el [objeto de aprendizaje](#) antes mencionado. Observa esta descripción mediante la opción 'enunciado'. Pasa luego a 'sugerencia'. Acaba con 'visualización'.

En \mathbf{R}^n , se pueden encontrar fácilmente cotas para esta función:

$$m = 0 \leq f(\mathbf{x}) \leq 0 - e^{-1} + 20 + e = 22.3504 = M,$$

ya que el primer término tiende a 0 cuando $x_i \rightarrow \infty$, y el segundo alcanza su mínimo cuando $\cos(2\pi x_i) = -1, \forall x_i$. Observa que esta cota es independiente de n .

Sin embargo, en $[-a, a]^n$, $a > 0$, se puede obtener la siguiente cota superior aproximada, más baja que la anterior, M , que considera la variabilidad de la primera exponencial. Este término alcanza su máximo para $x_i = |a|$ y vale

$$-20 \exp\left(-0.2 \sqrt{\frac{1}{n} \sum_{i=1}^n x_i^2}\right) \leq -20 \exp\left(-0.2 \sqrt{\frac{1}{n} n a^2}\right) = -20 \exp(-0.2a).$$

Sin embargo, hay que considerar la variabilidad del segundo término exponencial, que alcanza su máximo, $-e^{-1}$, en $x_i = k \pm 0.5$, siendo k un entero:

$$-\exp\left(\frac{1}{n} \sum_{i=1}^n \cos(2\pi x_i)\right) \leq -\exp\left(\frac{1}{n} (-n)\right) = -e^{-1}.$$

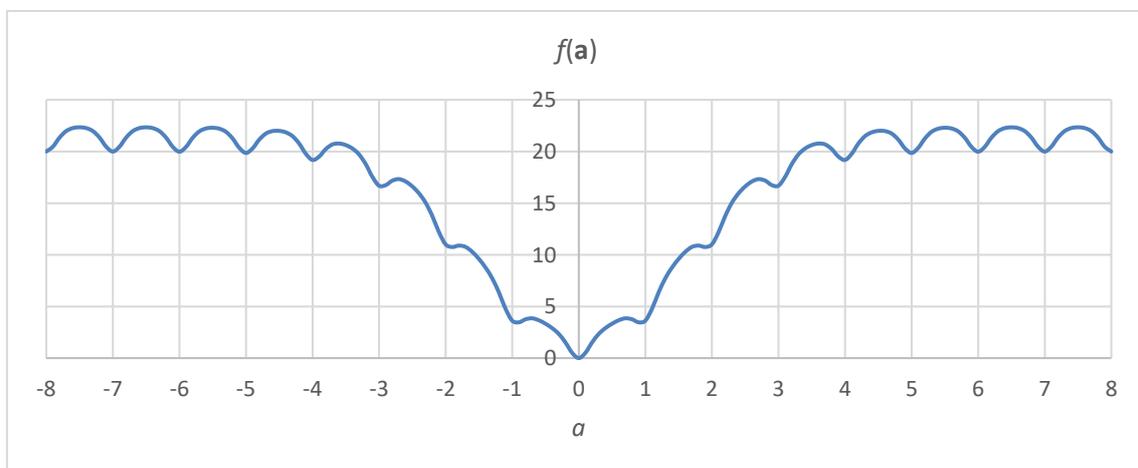


Figura 1. Gráfica de $f(a)$, para $a = (a, \dots, a)$ al variar a

Mira ahora la figura 1. Considerando que para $k \geq 4$ la variabilidad del primer término es menor que la del segundo, se puede tomar el valor de f en $\alpha = [a] + 0.5$, siendo $[a]$ el entero más próximo a a (observa que el segundo término vale e^{-1} en α), como una cota superior aproximada para f en $[-a, a]^n$:

$$0 \leq f(\mathbf{x}) \leq -20e^{-0.2\alpha} - e^{-1} + 20 + e = 20(1 - e^{-0.2\alpha}) + e - e^{-1} = M_\alpha.$$

La tabla siguiente presenta este valor para $a = 1, \dots, 10$.

a	1	2	3	4	5	6	7	8	9	10
M_α	7.534	10.220	12.419	14.219	15.693	16.900	17.888	18.696	19.360	19.901

Otro ejemplo: toma $a = 32.768$ (un valor habitual en la literatura para esta función); en tal caso $\alpha = 33.5$, y $M_{32.768} = 22.32$.

La figura 2 presenta la gráfica (para $n = 2$) de la función de Ackley en el dominio de búsqueda en que las variables están entre -5 y $+5$. Multimodal, ¿verdad?

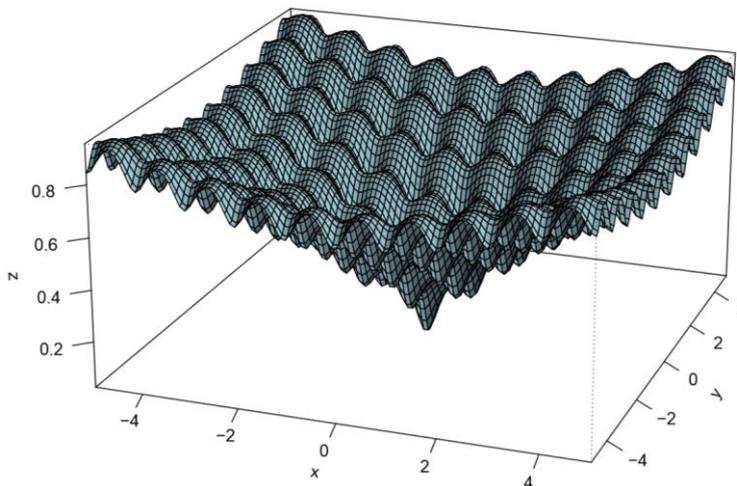


Figura 2. Función de Ackley para $n = 2$

Función de Griewank

El dominio de búsqueda habitual para esta función es $-600 \leq x_i \leq +600$. Su fórmula es:

$$f(\mathbf{x}) = \sum_{i=1}^n \frac{x_i^2}{4000} - \prod_{i=1}^n \cos\left(\frac{x_i}{\sqrt{i}}\right) + 1.$$

Visualizada en ese dominio, la función de Griewank tiene una sorpresa importante y parece no merecer estar en este apartado dedicado a funciones multimodales, como muestra la figura 3.

Te sugerimos que utilices de nuevo el [objeto de aprendizaje](#) antes mencionado para descubrir tal sorpresa. Cuando concluyas el estudio podrás comprobar que tiene varios mínimos locales regularmente distribuidos, como puedes observar en la figura 4, que muestra su gráfica restringida a $-10 \leq x_i \leq +10$, lo que refleja la complejidad de esta función.

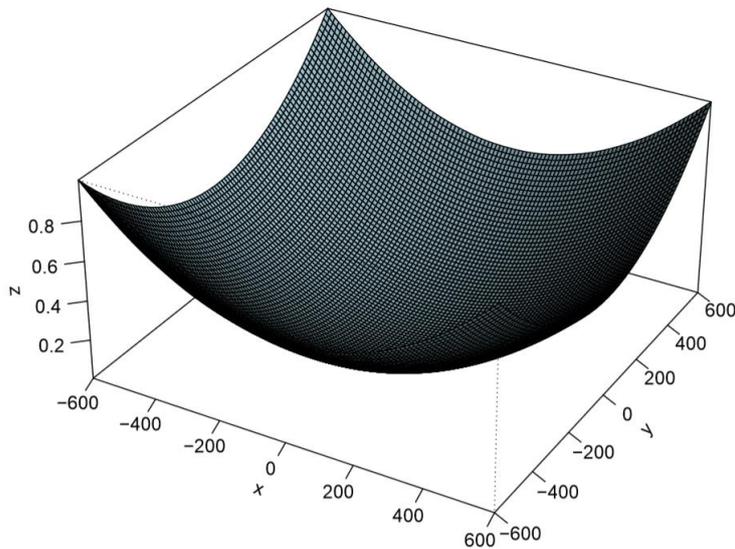


Figura 3. Función de Griewank para $n = 2$

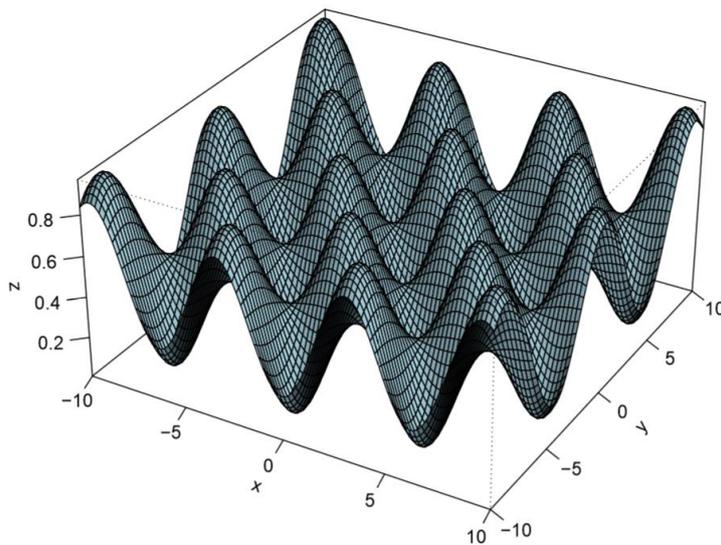


Figura 4. Función de Griewank para $n = 2$ restringida al dominio $-10 \leq x_i \leq +10$

Tiene, pues, muchos mínimos locales, pero su mínimo global es $m = 0$, que se alcanza en $\mathbf{x} = (0, \dots, 0)$. Los mínimos, sin embargo, están regularmente distribuidos.

En el dominio original, $-600 \leq x_i \leq +600$, $M = 90n + 1$, valor alcanzado cerca de $\mathbf{x} = (\pm 600, \dots, \pm 600)$ es un valor muy aproximado a su máximo absoluto (el producto de los cosenos es despreciable para n grande; por ejemplo menor que 10^{-5} para $n > 11$).

Función de Rastrigin

Se trata de una función típicamente no convexa que constituye un ejemplo tipo de función no lineal multimodal, caracterizada por un número muy grande de mínimos locales, aunque regularmente distribuidos, en su dominio habitual de definición. La función se define en un dominio n -dimensional, x_i variando entre -5.12 y $+5.12$:

$$f(\mathbf{x}) = 10n + \sum_{i=1}^n [x_i^2 - 10 \cos(2\pi x_i)].$$

Utiliza, una vez más, el [objeto de aprendizaje](#) para observar sus características. Son las siguientes.

Tiene un mínimo global, $m = 0$, en $\mathbf{x} = (0, \dots, 0)$.

Por otra parte, en el hipercubo $[-5.12, 5.12]^n$ esta función alcanza su máximo cuando cada uno de los sumandos (idénticos respecto de cada variable) de la función alcanza su máximo. Así, pues, considera la función

$$g(x) = x^2 - 10 \cos(2\pi x).$$

Se puede ver que esta función es par, y es 'globalmente' creciente, debido al término x^2 , pero con cierto carácter oscilatorio debido al término del coseno. Este coseno alcanza su máximo en $x = k \pm 0.5$, siendo k un entero. Es fácil ver que su máximo en $[0, 5.12]$ se alcanza en $x = 4.5$. Así, el máximo en $[-5.12, 5.12]$, alcanzado, por ejemplo, en $x = (4.5, \dots, 4.5)$ es:

$$M = 10n + n(4.5^2 + 10) = 40.25n.$$

Puedes ver su representación 3D en la figura 5. ¡Claramente multimodal!

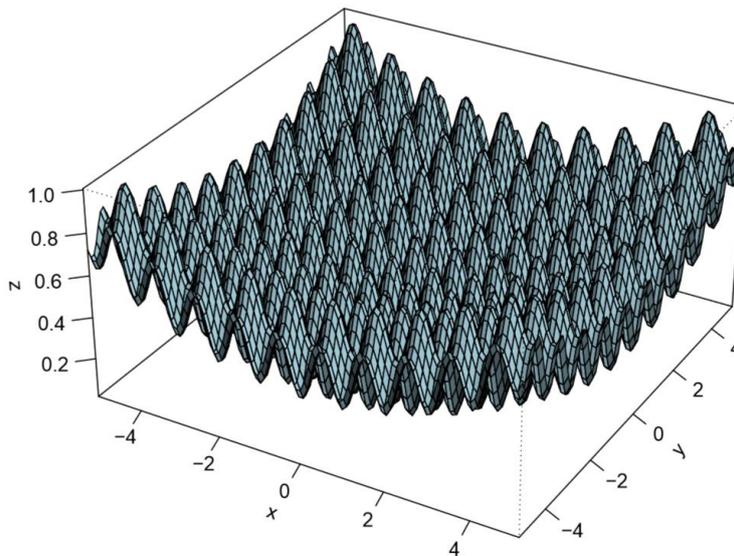


Figura 5. Función de Rastrigin para $n = 2$

Función de Schwefel

A veces denominada función 2.6 de Schwefel, es también una función multimodal típica. Como función de benchmarking se define habitualmente en un dominio compacto en el que las variables x_i varían entre -500 y $+500$. La función se expresa:

$$f(\mathbf{x}) = 418.982n - \sum_{i=1}^n x_i \sin(\sqrt{|x_i|}).$$

Su gráfica se ve en la figura 6.

Una vez más, utiliza [objeto de aprendizaje](#) para observar sus características.

Se sabe que la función tiene un mínimo global, $m = 0$, en $x = (420.9687, \dots, 420.9687)$. Típicamente, los algoritmos de optimización tienden a buscar en la dirección equivocada.

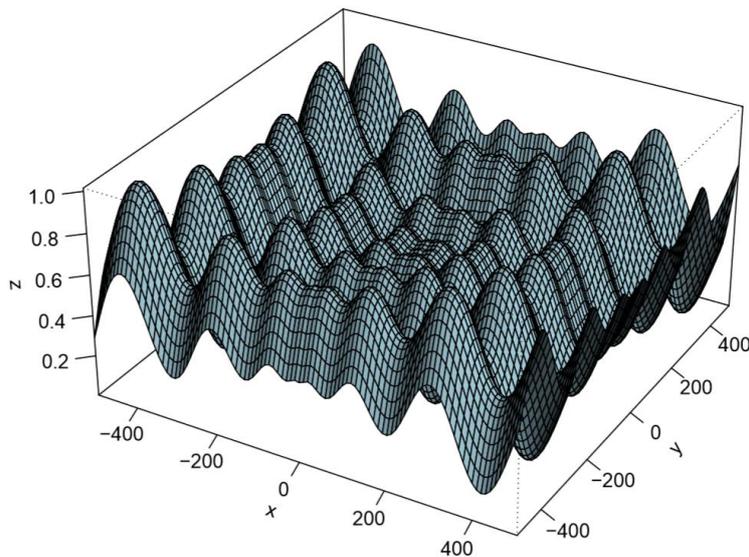


Figura 6. Función de Schwefel para $n = 2$

Obtener su máximo en $[-500, 500]^n$ es inmediato. Observa que la función $g(\mathbf{x}) = \sum_{i=1}^n x_i \sin(\sqrt{|x_i|})$ es impar en \mathbf{x} , es decir, $g(-\mathbf{x}) = -g(\mathbf{x})$. Por tanto, su máximo se obtiene en la ubicación simétrica (respecto al origen) de su mínimo. Así, el máximo de f es $M = 2 \times 418.982 n$ y se obtiene en $x = (-420.9687, \dots, -420.9687)$.

OTRAS FUNCIONES

Incluimos en este apartado algunos ejemplos de funciones cuya dificultad no estriba en la multimodalidad, sino en la dificultad que algoritmos tipo gradiente experimentan al enfrentarse a otras características, tales como vastas zonas con una variación muy pequeña de la función, es decir, zonas muy planas, valles muy alargados, etc. Las dos primeras son ejemplos de funciones que se caracterizan por tener su mínimo global en 'el fondo' de un valle muy alargado, donde la variación de la función es muy ligera. La convergencia al óptimo global es muy difícil. Por eso estas funciones son importantes en el benchmarking para algoritmos de tipo gradiente, que experimentan grandes dificultades para llegar a ese punto (¡o esos puntos!). Un regalo final cierra la colección.

Función de Dixon-Price

La función de Dixon-Price se evalúa habitualmente en el espacio de búsqueda definido por $-10 \leq x_i \leq +10$. La expresión para esta función es:

$$f(\mathbf{x}) = (x_1 - 1)^2 + \sum_{i=2}^n i(2x_i^2 - x_{i-1})^2,$$

y la figura 7 muestra una representación 3D. Te remitimos de nuevo a este [objeto de aprendizaje](#) para un estudio más detallado de esta función.

La función de Dixon-Price alcanza su mínimo global, $m = 0$, en los dos puntos dados por $x_i = 2^{-\frac{2^i-2}{2^i}}$ para $i = 1, 2, \dots, n-1$; $x_n = \pm 2^{-\frac{2^n-2}{2^n}}$. Te retamos a que compruebes esto analíticamente. Te sugerimos que empieces con valores bajos de n , hasta conseguir ver la generalización al caso general de n variables.

Observamos que esta función es una suma de números positivos. El primero alcanza su máximo, obviamente, cuando $x_1 = -10$, y los términos del sumatorio son mayores cuanto mayores sean los valores de x_i en valor absoluto, es decir, 10 o -10 , y menores sean los x_{i-1} , es decir -10 . Por tanto, esta función alcanza su máximo en $(-10, \dots, -10)$.

El máximo, M , es $121 + \sum_{i=2}^n i(200 + 10)^2 = 121 + 210^2 \sum_{i=2}^n i = 121 + 210^2 \frac{2+n}{2}(n-1)$.

Por ejemplo, para $n = 2$, se tiene $M = 121 + 2 \times 210^2 = 88321$.

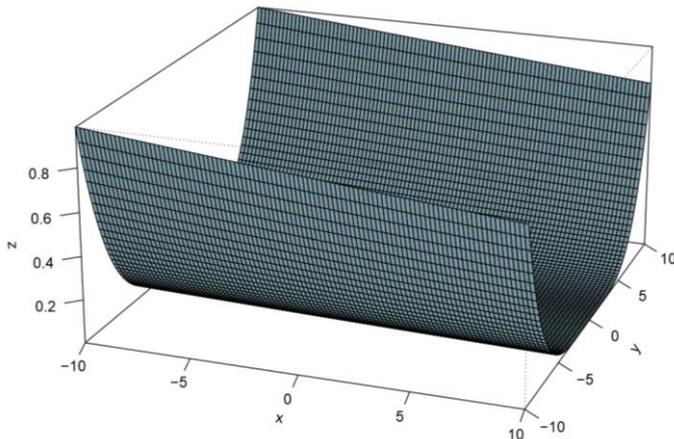


Figura 7. Función de Dixon-Price para $n = 2$

Función de Rosenbrock

La función de Rosenbrock, también conocida como función Valle o función Banana, es unimodal y alcanza su mínimo global en una región de tipo parabólico. Se define habitualmente en el dominio $-10 \leq x_i \leq +10$. Su expresión es:

$$f(\mathbf{x}) = \sum_{i=1}^{n-1} [100(x_{i+1} - x_i^2)^2 + (x_i - 1)^2].$$

Su representación 3D puede verse en la figura 8.

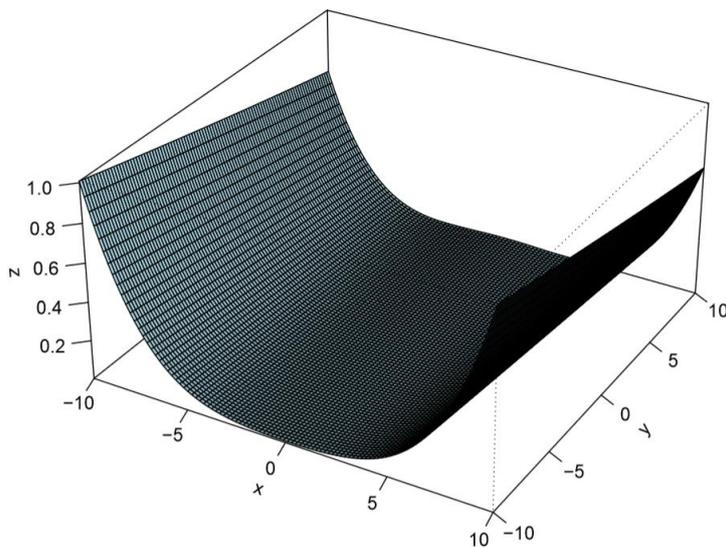


Figura 8. Función de Rosenbrock para $n = 2$

Puedes realizar un estudio más detallado utilizando el [objeto de aprendizaje](#).

El mínimo global, $m = 0$, lo alcanza en $\mathbf{x} = (1, \dots, 1)$ y su máximo en el dominio dado, $M = 1.2221 \times 10^6(n - 1)$, se alcanza en $\mathbf{x} = (-10, \dots, -10)$.

Función de Qing

Finalmente, la función de Qing es otra función multimodal camuflada. Se define habitualmente en el dominio determinado por $-500 \leq x_i \leq +500$. Esta función se expresa mediante:

$$f(\mathbf{x}) = \sum_{i=1}^n (x_i^2 - i)^2.$$

Puede verse su gráfica 3D en la figura 9. Sin embargo, esta gráfica es 'demasiado' alejada como para observar su naturaleza multimodal.

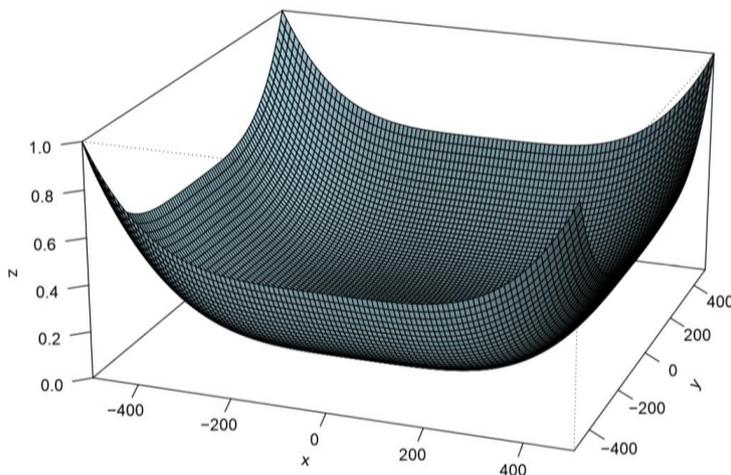


Figura 9. Función de Qing para $n = 2$

Es fácil ver que el mínimo global de esta función, $m = 0$, se alcanza en los 2^n puntos de la forma $\mathbf{x} = (\pm\sqrt{i})$. No dejes de utilizar cualquier representador gráfico de funciones del que dispongas, para comprobar esto para $n = 1$ (dos mínimos globales) y $n = 2$ (cuatro mínimos globales).

Su máximo, $M = \sum_{i=1}^n (2.5 \times 10^5 - i)^2$, se alcanza en $\mathbf{x} = (\pm 500, \dots, \pm 500)$.

5 Cierre

Te hemos presentado en este artículo un conjunto de funciones del benchmarking de optimización mono-objetivo. Además, te hemos proporcionado unas ideas simples que te permitan ser consciente de las sorpresas variadas que depara la optimización. Te hemos propuesto la utilización de un objeto de aprendizaje interactivo que te ayudará a adquirir esta competencia básica. Con frecuencia, esta competencia no es proporcionada en los cursos elementales de Matemáticas. Sin embargo, es un hecho que gran parte de los problemas de optimización reales son altamente complejos, y resulta esencial que los estudiantes de ingeniería, tengáis conciencia de las dificultades que entraña la optimización. Así, que tengas una visión amplia de la naturaleza del problema que como ingeniero tengas que resolver alguna vez, es una

competencia a largo plazo de este artículo. El elemento multidisciplinar que te presentamos trata, finalmente, pues, de transmitirte competencias transversales, como pensamiento crítico, conocimiento de problemas contemporáneos y análisis y resolución de problemas.

6 Bibliografía

- Bossek, J. (2017). smooF: Single- and Multi-Objective Optimization Test Functions. *The R Journal*, 9 (1), 103–113.
- Izquierdo, J. (2016). Optimización en el mundo real: la multimodalidad. <http://hdl.handle.net/10251/68322>.
- Jamil, M., Yang, X.-S. (2013). A Literature Survey of Benchmark Functions For Global Optimization Problems. *International Journal of Mathematical Modelling and Numerical Optimisation*, 4 (2), 150–194.
- Surjanovic, S., Bingham, D. (2017). Virtual library of Simulation Experiments: Test functions and data set. <http://www.sfu.ca/~ssurjano/optimization.html>.
- R Development Core Team (2011). A language and environment for statistical computing. Vienna: R Foundation for Statistical Computing. ISBN 3-900051-07-0.
- GAMS World, GLOBAL Library, Disponible Online en: <http://www.gamsworld.org/global/globallib.html>.
- GO Test Problems, Disponible Online en: http://www-optima.amp.i.kyoto-u.ac.jp/member/student/hedar/Hedar_files/TestGO.htm.
- Gould, N. I. M.; Orban, D. y Toint, P. L. "CUTEr, A Constrained and Un-constrained Testing Environment, Revisited," Disponible Online en: <http://cuter.rl.ac.uk/cuter-www/problems.html>.
- Hedar, A.-R. Global Optimization Test Problems. Disponible Online en: http://www-optima.amp.i.kyoto-u.ac.jp/member/student/hedar/Hedar_files/TestGO.htm.
- Izquierdo, J. (2009). Funciones Benchmarking Optimización. Disponible Online en: <https://laboratoriosvirtuales.upv.es/eslabon/P/>.