ALGORITMO BFGS

Grupo 1

Miembros:

Zakaria Lasry Sahraoui

Daniel Acosta Luna

Ismael Tse Perdomo Rodriguez

8 de diciembre de 2025

Explicación algoritmo BFGS

En problemas no lineales complejos se requiere un enfoque iterativo robusto para minimizar la función de coste. El algoritmo BFGS, desarrollado independientemente por sus cuatro autores en 1970, fue diseñado para superar las limitaciones computacionales y teóricas del método de Newton clásico. El método de Newton basa su eficacia en aproximar localmente la función objetivo en un punto actual x_k mediante una función cuadrática derivada de la expansión de Taylor de segundo orden, expresada como $m_k(p) = f(x_k) + \nabla f(x_k)^T p + \frac{1}{2} p^T B_k p$. Esta aproximación depende crucialmente de la matriz Hessiana B_k, que contiene las segundas derivadas y define la curvatura de la función. El paso ideal de Newton se obtiene minimizando esta función cuadrática, lo que resulta en la dirección $p_k = -B_k^{-1} \nabla f(x_k)$. Sin embargo, el cálculo de la Hessiana es computacionalmente costoso, creciendo cuadráticamente con la dimensión del problema, y en ocasiones la matriz puede no estar disponible analíticamente. Más crítico aún es el requisito de que B_k sea definida positiva; si no lo es, la aproximación cuadrática podría dirigir la búsqueda hacia un máximo o un punto de silla en lugar de un mínimo, haciendo fallar al algoritmo.

Para solventar estos problemas, el método BFGS propone trabajar con una aproximación de la matriz Hessiana en lugar de calcular la exacta. Esta matriz aproximada se construye iterativamente utilizando únicamente la información del gradiente, bajo la premisa de que la Hessiana no debería variar drásticamente entre pasos consecutivos. La construcción de esta matriz debe satisfacer dos restricciones fundamentales: primero, debe ser simétrica; segundo, debe cumplir la "condición de la secante". Esta condición garantiza que el gradiente de la aproximación cuadrática coincida con el gradiente de la función original en el paso anterior. Matemáticamente, si definimos el desplazamiento como $s_k = x_{k+1} - x_k$ y el cambio en el gradiente como $y_k = \nabla f(x_{k+1}) - \nabla f(x_k)$, la condición se formula como $B_{k+1} s_k = y_k$. Esto asegura que la curvatura estimada sea consistente con el cambio observado en la pendiente de la función entre iteraciones.

Aun satisfaciendo estas condiciones, existe una familia de posibles matrices de actualización, y este algoritmo se diferencia en cómo elige la "mejor" matriz que sea lo más cercana posible a la aproximación de la iteración anterior. El BFGS minimiza la norma de Frobenius ponderada de la diferencia entre las inversas de las matrices Hessianas, buscando resolver el problema de optimización $\min ||H_{k+1} - H_k||_W$, donde H_k es la aproximación de la inversa de la Hessiana. Esta formulación le otorga una ventaja crítica: su capacidad para preservar la propiedad de definición positiva. Si la aproximación inicial H_0 es definida positiva (comúnmente la matriz identidad), el algoritmo garantiza que todas las actualizaciones subsiguientes mantendrán esta propiedad, asegurando así que cada paso de búsqueda sea una dirección de descenso válida. Finalmente, la eficiencia del BFGS radica en que la actualización de H_{k+1} se realiza mediante fórmulas explícitas que involucran solo operaciones de álgebra lineal de bajo coste, evitando la costosa inversión de matrices $O(n^3)$ en cada paso.