Trabalho de Otimização Numérica

Margarida Biscaia Caleiras
Mestrado em Matemática, Universidade de Coimbra

Pretende-se aproximar uma curva $y = f(x)$, $x \in [a,b]$, por uma função g com as seguintes características

(1) é linear por ramos e tem apenas um ponto de quebra u,

(2) é contínua,

(3) $g(a) = f(a)$ e $g(b) = f(b)$,

(4) g minimiza o erro $\int_a^b (f(x) - g(x))^2 dx$.

Vamos considerar o caso de estudo $f(x) = (x-1)^2$ e $[a,b] = [0,3]$.

Nota: Os exercícios de matlab estão no ficheiro **Matlab.m** devidamente identificados.

Exercício 1. O problema de otimização (P) a resolver é

$$\min_{u,v\in\mathbb{R}} \int_a^b (f(x) - g(x))^2 dx$$

pelo que as variáveis de decisão são u e v, onde $v = g(u)$.

Para representar o problema graficamente, vamos começar por determinar a equação de g.

Consideremos

$$g(x) = \begin{cases} g_1(x), & x \in [a,u] \\ g_2(x), & x \in (u,b] \end{cases}$$

Como g linear por ramos vem $g_1(x) = m_1 x + c_1$ e $g_2(x) = m_2 x + c_2$. Além disso, g contínua pelo que $g_1(u) = g_2(u) = v$.

Como $(a, f(a))$ e (u,v) pontos de g_1 temos

$$m_1 = \frac{f(a) - v}{a - u}, \quad c_1 = f(a) - \frac{f(a) - v}{a - u} \times a.$$

Assim,

$$\begin{aligned} g_1(x) &= \frac{f(a) - v}{a - u} \times x + f(a) - \frac{f(a) - v}{a - u} \times a \\ &= \frac{f(a) - v}{a - u}(x - a) + f(a). \end{aligned}$$

Como $(b, f(b))$ e (u,v) pontos de g_2 temos

$$m_2 = \frac{f(b) - v}{b - u}, \quad c_2 = f(b) - \frac{f(b) - v}{b - u} \times b.$$

Assim,

$$\begin{aligned} g_2(x) &= \frac{f(b)-v}{b-u} \times x + f(b) - \frac{f(b)-v}{b-u} \times b \\ &= \frac{f(b)-v}{b-u}(x-b) + f(b). \end{aligned}$$

Obtivemos a seguinte representação para um ponto (u, v) aleatoriamente escolhido.

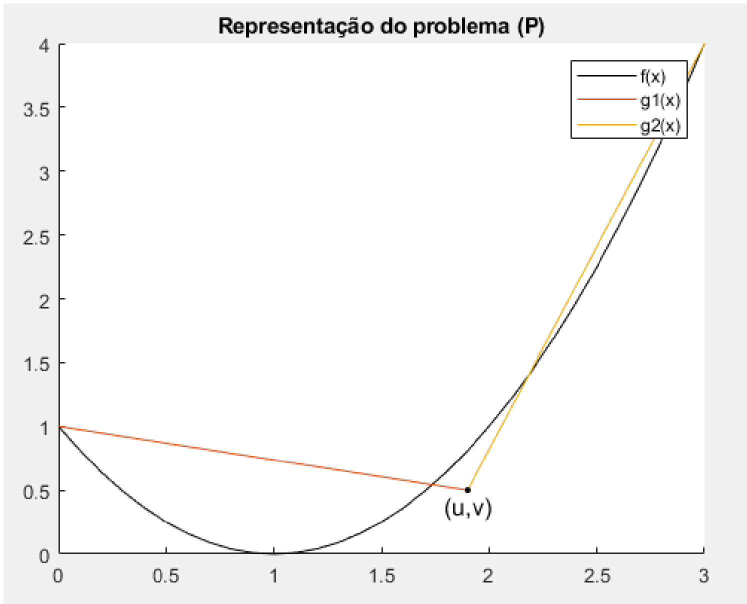

Figura 1: Exemplo gráfico do modelo (P)

Geometricamente, o problema (P) consiste em encontrar o ponto de quebra (u, v) que minimiza a área assinalada na figura 2.

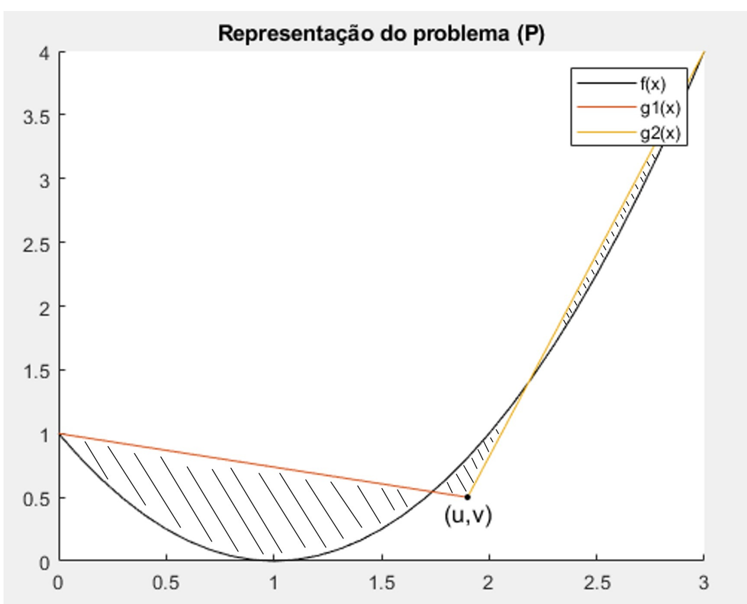

Figura 2: Exemplo gráfico do modelo (P)

Exercício 2. A função objetivo de (P) é uma norma no espaço das funções pelo que é uma função convexa. Assim, se existir minimizante local, ele é global. Além disso, a função está definida num intervalo fechado $[a, b]$ onde é contínua (pois f e g contínuas), logo pelo teorema de Weirstrass existe minimizante global. Assim, devido à convexidade do problema, (P) tem solução e todo o minimizante local é global.

Exercício 3. Para encontrar a solução ótima do problema (P) foram utilizados dois métodos de otimização: fminsearch (que implementa o método simplex Nelder-Mead) e fminunc (que implementa o método Quasi-Newton). Cada um deles vai devolver a solução ótima, xOpt, e o valor ótimo, fOpt, assim como o número de iterações (iterations), o número de vezes que a função objetivo foi avaliada (funcCount) e a precisão do resultado (firstorderopt). Obtivemos os resultados da tabela 1. Como seria de esperar, o método de Nelder-Mead obtém a solução ótima em muito mais iterações por ser um método de procura direta sem utilização de derivadas. O método Quasi-Newton é um método de procura unidirecional e converge mais rapidamente. Na figura 3, temos a representação gráfica da solução obtida pelos dois métodos.

Comparação entre os dois métodos		
	Nelder-Mead	Quasi-Newton
Número de iterações	54	7
Número de avaliações	104	27
Precisão do resultado	-	4.8180×10^{-7}
Solução ótima	$(1.5000, -0.3125)$	$(1.5000, -0.3125)$
Valor ótimo	0.1898	0.1898

Tabela 1: Resultados dos dois métodos utilizados para (P)

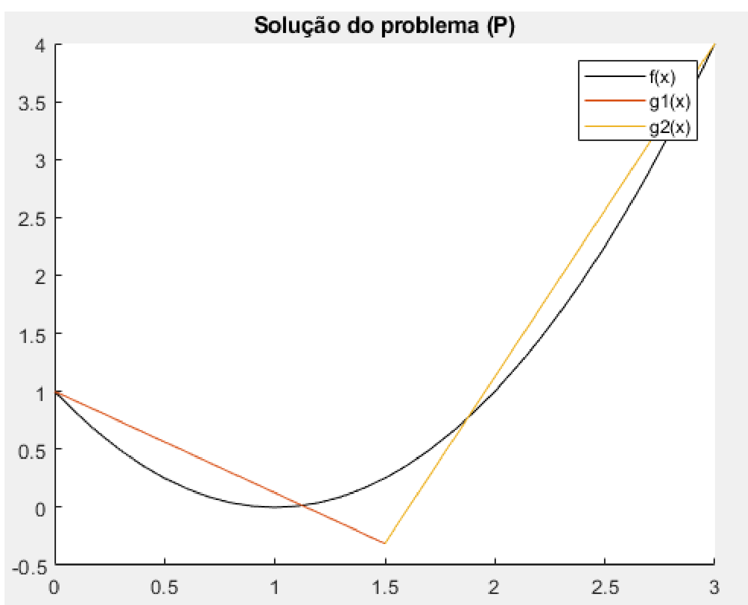

Figura 3: Solução ótima do problema (P)

Exercício 4. Temos

$$\begin{aligned}
\int_a^b (f(x)-g(x))^2 dx &= \int_a^u (f(x)-g_1(x))^2 dx + \int_u^b (f(x)-g_2(x))^2 dx \\
&= \int_a^u \left[(x-1)^2 - \left(\frac{f(a)-v}{a-u}(x-a)+f(a)\right)\right]^2 dx + \\
&\quad + \int_u^b \left[(x-1)^2 - \left(\frac{f(b)-v}{b-u}(x-b)+f(b)\right)\right]^2 dx \\
&= -\frac{1}{30}(a-u)(-5f(a)(3a^2+2a(u-4)+(u-4)u-2v+6) - \\
&- 5v(a^2+2a(u-2)+u(3u-8)+6) + \\
&+ 6(a^4+a^3(u-5)+a^2((u-5)u+10)+a(u((u-5)u+10)-10) + \\
&+ u(u((u-5)u+10)-10)+5)+10f(a)^2+10v^2) + \\
&+ \frac{1}{30}(b-u)(-5f(b)(3b^2+2b(u-4)+(u-4)u-2v+6) - \\
&- 5v(b^2+2b(u-2)+u(3u-8)+6) + \\
&+ 6(b^4+b^3(u-5)+b^2((u-5)u+10)+b(u((u-5)u+10)-10) + \\
&+ u(u((u-5)u+10)-10)+5)+10f(b)^2+10v^2)
\end{aligned}$$

Cálculo auxiliar: [1],[2]

Seja $h(u,v) = \int_a^b (f(x)-g(x))^2 dx$. Temos

$$\begin{aligned}
\frac{\partial h}{\partial u}(u,v) &= \frac{1}{6}(-f(a)(a^2+2a(u-2)+3u^2-8u-2v+6)+a^2v+2f(a)^2 + \\
&+ 2u((a+8)v-12)-4av+6u^4-24u^3-9u^2(v-4)+2v^2-6v+6) + \\
&+ \frac{1}{6}(f(b)(b^2+2b(u-2)+3u^2-8u-2v+6)-b^2v-2f(b)^2-2(b+8)uv + \\
&+ 4bv-6u^4+24u^3+9u^2(v-4)+24u-2v^2+6v-6)
\end{aligned}$$

Cálculo auxiliar: [3],[4]

$$\begin{aligned}
\frac{\partial h}{\partial v}(u,v) &= \frac{1}{6}(a-u)(a^2-2f(a)+2a(u-2)+3u^2-8u-4v+6) - \\
&- \frac{1}{6}(b-u)(b^2-2f(b)+2b(u-2)+3u^2-8u-4v+6)
\end{aligned}$$

Cálculo auxiliar: [5],[6]

Assim, para $(\bar{u},\bar{v}) = (1.5, -0.3125)$ obtidos no exercício 3 e para os valores $a, f(a), b, f(b)$ dados no enunciado vem

$$\frac{\partial h}{\partial u}(\bar{u},\bar{v}) = \frac{\partial h}{\partial v}(\bar{u},\bar{v}) = 0$$

pelo que $(\bar{u},\bar{v})$ satisfaz as condições de otimalidade de primeira ordem e é candidato a minimizante local de h.

Cálculo auxiliar: ficheiro matlab calculos.m

Além disso,

$$\frac{\partial^2 h}{\partial u^2}(u,v) = \frac{1}{3}(-f(a)(a+3u-4)+av+12u^3-36u^2-9u(v-4)+8v-12)+ \\ +\frac{1}{3}(f(b)(b+3u-4)-bv-12u^3+36u^2+9u(v-4)-8v+12)$$

Cálculo auxiliar: [7],[8]

$$\frac{\partial^2 h}{\partial v^2}(u,v) = -\frac{2}{3}(a-u)+\frac{2}{3}(b-u)$$

Cálculo auxiliar: [9],[10]

$$\frac{\partial^2 h}{\partial u \partial v}(u,v) = \frac{1}{6}((a-b)(a+b+2u-4)+2f(a)-2f(b))$$

Cálculo auxiliar: [11]

Pelo que,

$$\frac{\partial^2 h}{\partial u^2}(\bar{u},\bar{v}) = 4.8125, \quad \frac{\partial^2 h}{\partial v^2}(\bar{u},\bar{v}) = 2, \quad \frac{\partial^2 h}{\partial u \partial v}(\bar{u},\bar{v}) = -2$$

Cálculo auxiliar: ficheiro matlab calculos.m

Como a matriz

$$\begin{bmatrix} \frac{\partial^2 h}{\partial u^2} & \frac{\partial^2 h}{\partial u \partial v} \\ \\ \frac{\partial^2 h}{\partial u \partial v} & \frac{\partial^2 h}{\partial v^2} \end{bmatrix}_{(\bar{u},\bar{v})} = \begin{bmatrix} 4.8125 & -2 \\ -2 & 2 \end{bmatrix}$$

é positiva definida, então $(\bar{u},\bar{v})$ satisfaz as condições suficientes de 2ª ordem, pelo que $(\bar{u},\bar{v})$ é minimizante local de h (e consequentemente minimizante global, pelo exercício 2).

Exercício 5. O problema de otimização (P2) a resolver é

$$\begin{aligned} &\min_{u,v\in\mathbb{R}} \int_a^b (f(x)-g(x))^2 dx \\ &\text{s.a} \quad g(u) \geq f(u) \end{aligned}$$

Provámos no exercício 2 que a função objetivo do problema é convexa.

A restrição $c(u,v) = g(u) - f(u) \geq 0$ é convexa se $-c(u,v) = f(u) - g(u)$ for convexa. Temos,

$$\bigtriangledown(-c(u,v)) = \begin{bmatrix} f'(u) \\ -1 \end{bmatrix}, \quad \bigtriangledown^2(-c(u,v)) = \begin{bmatrix} f''(u) & 0 \\ 0 & 0 \end{bmatrix}$$

pelo que $\bigtriangledown^2(-c(u,v))$ é semi definida positiva quando f convexa (pois nesse caso $f''(u) \geq 0$).

No nosso exemplo, $f(x) = (x-1)^2$ é uma parábola com côncavidade voltada para cima pelo que f é convexa. Então $f(u) - g(u)$ convexa e temos um problema convexo.

Exercício 6. Para encontrar a solução ótima do problema (P2) foram utilizados dois métodos de otimização com o comando fmincon: o método do ponto interior e o método do conjunto ativo, ambos métodos numéricos para otimização não linear com restrições. Cada um deles vai devolver a solução ótima, xOpt, e o valor ótimo, fOpt, assim como o número de iterações (iterations), o número de vezes que a função objetivo foi avaliada (funcCount) e a precisão do resultado (firstorderopt). Obtivemos os resultados da tabela 2. Como podemos observar, o método do conjunto ativo converge em menos iterações e tem uma melhor precisão. Na figura 4, temos a representação gráfica da solução obtida pelos dois métodos.

Comparação entre os dois métodos		
	Ponto interior	Conjunto ativo
Número de iterações	8	6
Número de avaliações	29	22
Precisão do resultado	7.1776×10^{-8}	3.9259×10^{-7}
Solução ótima	$(1.5000, 0.2500)$	$(1.5000, 0.2500)$
Valor ótimo	0.5063	0.5062

Tabela 2: Resultados dos dois métodos utilizados para (P2)

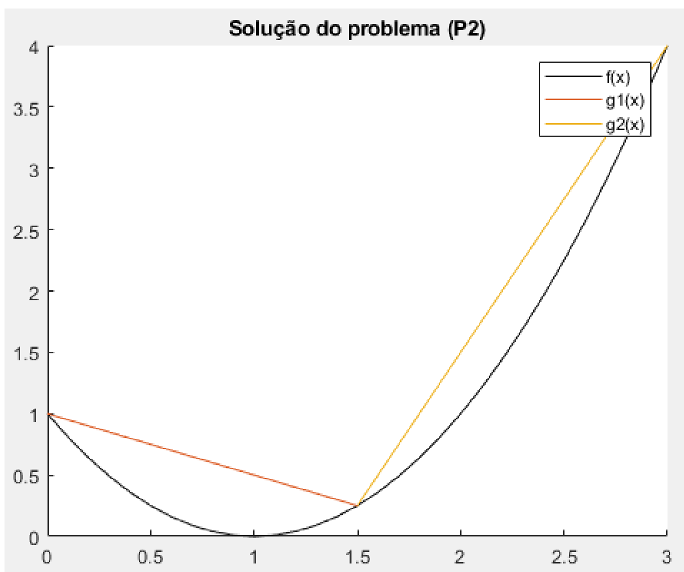

Figura 4: Solução ótima do problema (P2)

Exercício 7. Temos, para cada $\lambda \in [-2, 2]$,

$$\mathscr{L}(x, \lambda) = \int_a^b (f(x) - g(x))^2 dx \ - \lambda(g(u) - f(u)).$$

Assim, a função dual é dada por

$$L_d(\lambda) = \inf_{x \in \mathbb{R}^2} \mathscr{L}(x, \lambda)$$

Para obter a solução ótima deste problema utilizei a função fminsearch. Na figura 5, temos a representação da função dual em $\lambda \in [-2, 2]$.

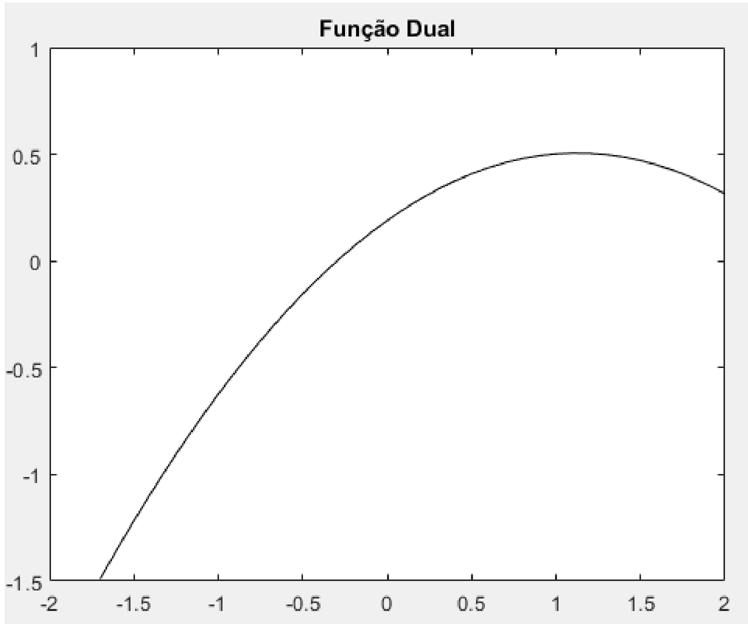

Figura 5: Função Dual

Exercício 8. A solução ótima do dual é $(1.1000,\ 0.5061)$.

Provámos no exercício 5 que o problema (P2) é convexo, logo o gap dual é 0.

Observação: Podemos facilmente confirmar que o gap dual é 0. No exercício 6 obtivemos o valor ótimo do problema (P2), que corresponde à solução ótima do problema primal. Como podemos observar na tabela 2, o valor ótimo coincide, quase certamente, com a solução do dual, pelo que o gap dual é 0.

Referências

[1] https://www.wolframalpha.com/input?i2d=true&i=Integrate%5B%5C%2840%29Power%5B%5C%2840%29Power%5B%5C%2840%29x-1%5C%2841%29%2C+2%5D-+%5C%2840%29Divide%5Bf%5C%2840%29a%5C%2841%29-v%2Ca-u%5D%5C%2840%29x-a%5C%2841%29+%2Bf%5C%2840%29a%5C%2841%29%5C%2841%29%5C%2841%29%2C2%5D%5C%2841%29%2C%7Bx%2Ca%2Cu%7D%5D

[2] https://www.wolframalpha.com/input?i2d=true&i=Integrate%5B%5C%2840%29Power%5B%5C%2840%29Power%5B%5C%2840%29x-1%5C%2841%29%2C+2%5D-+%5C%2840%29Divide%5Bf%5C%2840%29b%5C%2841%29-v%2Cb-u%5D%5C%2840%29x-b%5C%2841%29+%2Bf%5C%2840%29b%5C%2841%29%5C%2841%29%5C%2841%29%2C2%5D%5C%2841%29%2C%7Bx%2Cu%2Cb%7D%5D

[3] https://www.wolframalpha.com/input?i2d=true&i=D%5BIntegrate%5B%5C%2840%29Power%5B%5C%2840%29Power%5B%5C%2840%29x-1%5C%2841%29%2C+2%5D-+%5C%2840%29Divide%5Bf%5C%2840%29a%5C%2841%29-v%2Ca-u%5D%5C%2840%29x-a%5C%2841%29+%2Bf%5C%2840%29a%5C%2841%29%5C%2841%29%5C%2841%29%2C2%5D%5C%2841%29%2C%7Bx%2Ca%2Cu%7D%5D%2Cu%5D

[4] https://www.wolframalpha.com/input?i2d=true&i=D%5BIntegrate%5B%5C%2840%29Power%5B%5C%2840%29Power%5B%5C%2840%29x-1%5C%2841%29%2C+2%5D-+%5C%2840%29Divide%5Bf%5C%2840%29b%5C%2841%29-v%2Cb-u%5D%5C%2840%29x-b%5C%2841%29+%2Bf%5C%2840%29b%5C%2841%29%5C%2841%29%5C%2841%29%2C2%5D%5C%2841%29%2C%7Bx%2Cu%2Cb%7D%5D%2Cu%5D

[5] https://www.wolframalpha.com/input?i2d=true&i=D%5BIntegrate%5B%5C%2840%29Power%5B%5C%2840%29Power%5B%5C%2840%29x-1%5C%2841%29%2C+2%5D-+%5C%2840%29Divide%5Bf%5C%2840%29a%5C%2841%29-v%2Ca-u%5D%5C%2840%29x-a%5C%2841%29+%2Bf%5C%2840%29a%5C%2841%29%5C%2841%29%5C%2841%29%2C2%5D%5C%2841%29%2C%7Bx%2Ca%2Cu%7D%5D%2Cv%5D

[6] https://www.wolframalpha.com/input?i2d=true&i=D%5BIntegrate%5B%5C%2840%29Power%5B%5C%2840%29Power%5B%5C%2840%29x-1%5C%2841%29%2C+2%5D-+%5C%2840%29Divide%5Bf%5C%2840%29b%5C%2841%29-v%2Cb-u%5D%5C%2840%29x-b%5C%2841%29+%2Bf%5C%2840%29b%5C%2841%29%5C%2841%29%5C%2841%29%2C2%5D%5C%2841%29%2C%7Bx%2Cu%2Cb%7D%5D%2Cv%5D

[7] https://www.wolframalpha.com/input?i2d=true&i=++D%5BIntegrate%5B%5C%2840%29Power%5B%5C%2840%29Power%5B%5C%2840%29x-1%5C%2841%29%2C+2%5D-+%5C%2840%29Divide%5Bf%5C%2840%29a%5C%2841%29-v%2Ca-u%5D%5C%2840%29x-a%5C%2841%29+%2Bf%5C%2840%29a%5C%2841%29%5C%2841%29%5C%2841%29%2C2%5D%5C%2841%29%2C%7Bx%2Ca%2Cu%7D%5D%2C%7Bu%2C2%7D%5D

[8] https://www.wolframalpha.com/input?i2d=true&i=++D%5BIntegrate%5B%5C%2840%29Power%5B%5C%2840%29Power%5B%5C%2840%29x-1%5C%2841%29%2C+2%5D-+%5C%2840%29Divide%5Bf%5C%2840%29b%5C%2841%29-v%2Cb-u%5D%5C%2840%29x-b%5C%2841%29+%2Bf%5C%2840%29b%5C%2841%29%5C%2841%29%5C%2841%29%2C2%5D%5C%2841%29%2C%7Bx%2Cu%2Cb%7D%5D%2C%7Bu%2C2%7D%5D

[9] https://www.wolframalpha.com/input?i2d=true&i=++D%5BIntegrate%5B%5C%2840%29Power%5B%5C%2840%29Power%5B%5C%2840%29x-1%5C%2841%29%2C+2%5D-+%5C%2840%29Divide%5Bf%5C%2840%29a%5C%2841%29-v%2Ca-u%5D%5C%2840%29x-a%5C%2841%29+%2Bf%5C%2840%29a%5C%2841%29%5C%2841%29%5C%2841%29%2C2%5D%5C%2841%29%2C%7Bx%2Ca%2Cu%7D%5D%2C%7Bv%2C2%7D%5D

[10] https://www.wolframalpha.com/input?i2d=true&i=++D%5BIntegrate%5B%5C%2840%29Power%5B%5C%2840%29Power%5B%5C%2840%29x-1%5C%2841%29%2C+2%5D-+%5C%2840%29Divide%5Bf%5C%2840%29b%5C%2841%29-v%2Cb-u%5D%5C%2840%29x-b%5C%2841%29+%2Bf%5C%2840%29b%5C%2841%29%5C%2841%29%5C%2841%29%2C2%5D%5C%2841%29%2C%7Bx%2Cu%2Cb%7D%5D%2C%7Bv%2C2%7D%5D

[11] https://www.wolframalpha.com/input?i2d=true&i=Divide%5Bd%2Cdu%5DDivide%5B1%2C6%5D%5C%2840%29a-u%5C%2841%29%5C%2840%29Power%5Ba%2C2%5D-2f%5C%2840%29a%5C%2841%29%2B2a%5C%2840%29u-2%5C%2841%29%2B3Power%5Bu%2C2%5D-8u-4v%2B6%5C%2841%29-Divide%5B1%2C6%5D%5C%2840%29b-u%5C%2841%29%5C%2840%29Power%5Bb%2C2%5D-2f%5C%2840%29b%5C%2841%29%2B2b%5C%2840%29u-2%5C%2841%29%2B3Power%5Bu%2C2%5D-8u-4v%2B6%5C%2841%29