Entregable 3: Dualidad

14 de octubre de 2020

Ejercicio 1 - Interpretación geométrica de dualidad. 5.22 [1,p.280]

Considere el problema primario

$$(P): \quad \min_{x\in\mathbb{R}} f(x)$$
$$\text{sujeto a: } \; g(x) \leq 0.$$

con $f(x) = x$ y $g(x) = x^2 - 1$.

a) Resuelva *(P)* hallando el ínfimo f^* y la variable primal óptima x^*, de existir.

b) Obtenga y resuelva problema dual, hallando el supremo d^* y la variable dual óptima μ^*, de existir.

c) Analice si se cumple la condición de Slater y determine si se cumplen o no dualidad débil y fuerte.

d) Represente los conjuntos $\mathcal{S}$ y $\mathcal{A}$ junto a la recta de pendiente $-\mu^*$

$$\mathcal{S} = \{(t, z): \; z = f(x), \; t = g(x), \; x \in X\} \tag{1}$$
$$\mathcal{A} = \{(t, z): \; f(x) \leq z, \; g(x) \leq t, \; x \in X\} \tag{2}$$

e) Represente f^* y d^* y la condición de Slater en el esquema anterior, y compruebe geométricamente los resultados obtenidos en la parte c).

f) Repita las partes anteriores para $f(x) = x$ y $g(x) = x^2$.

g) Repita para $f(x) = x$ y $g(x) = |x|$. Observe que aquí el óptimo μ^* no es único, por lo que existen varías rectas de pendiente $-\mu^*$.

h) Repita para $f(x) = x^3$ y $g(x) = -x + 1$. (Nota: Dado que $\lim_{x\to-\infty} x^3 = -\infty$, el Lagrangiano no se minimiza anulando su gradiente.)

Ejercicio 2 - Minimización cuadrática con restricciones de caja

Sea el siguiente problema (QP) con parámetros α, β, y γ.

$$
\begin{aligned}
\min_{(x,y,z)\in\mathbb{R}^3} \ & \alpha_1 x^2 + \alpha_2 y^2 + \alpha_3 z^2 + \beta_1 xy + \beta_2 xz + \beta_3 yz + \gamma_1 x + \gamma_2 y + \gamma_3 z \\
\text{s. to: } & x \in [0,1] \\
& y \in [0,1] \\
& z \in [0,1]
\end{aligned}
$$

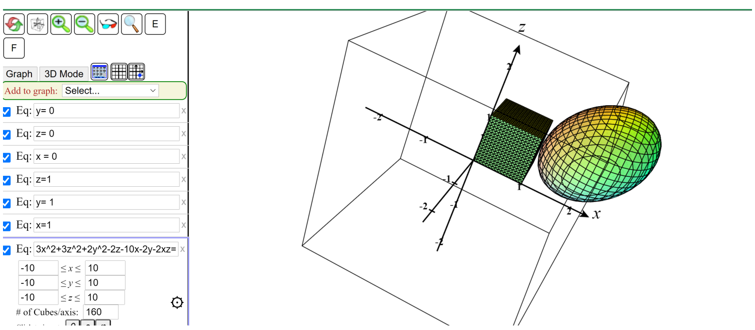
Plantear las condiciones de optimalidad de Karush Kuhn Tucker e interpretar geométricamente complementary slackness, asumiendo que la función objetivo es estrictamente convexa. En particular, justificar las siguientes afirmaciones

- Si se pone como hipótesis que el mínimo se da en el interior del cubo, entonces todos los multiplicadores son nulos y el gradiente se anula. Esto determina un punto. La hipótesis se descarta si este punto no pertenece al interior del cubo.
- Si se pone como hipótesis que el mínimo se da en el interior de una cara, entonces al menos cinco de los multiplicadores se anulan. Luego dos entradas del gradiente se anulan, y esto resulta en dos ecuaciones que determinan la posición de un punto en el plano de la cara. La hipótesis se descarta si este punto no pertenece al interior de la cara.
- Si se pone como hipótesis que el mínimo se da en el interior de una artista, entonces al menos cuatro de los multiplicadores se anulan. Luego una entrada del gradiente se anula, y esto resulta en una ecuación que determina la posición de un punto en la recta que contiene a la arista. La hipótesis se acepta si este punto no pertenece al interior de la arista.
- Si todas las hipótesis anteriores se descartan, entonces el mínimo se da en un vértice del cubo.

Usando ese procedimiento, obtenga analíticamente el mínimo del siguiente problema (QP).

$$
\begin{aligned}
\min_{(x,y,z)\in\mathbb{R}^3} \ & 3x^2 + 2y^2 + 3z^2 - 2xz - 10x - 2y - 2z \\
\text{s. to: } & x \in [0,1] \\
& y \in [0,1] \\
& z \in [0,1]
\end{aligned}
$$

La siguiente herramienta de representación gráfica ayuda a interpretar el resultado.
`https://www.monroecc.edu/faculty/paulseeburger/calcnsf/CalcPlot3D/`



Ejercicio 3 - *Dualidad en supported vector machines (SVMs)*

Considere la siguiente versión de SVM [1,p.427]: Dados $x_j \in \mathbb{R}^n$ y $y_j \in \{-1, 1\}$, $j = 1, \ldots, m$, se busca el hiperplano separador $a^T x + b = 0$ cuyos parámetros $a \in \mathbb{R}^n$ y $b \in \mathbb{R}$ optimizan

$$(a^*, b^*, \epsilon^*) = \arg \min_{(a,b,\epsilon)} \frac{1}{2}\|a\|^2 \qquad (3)$$
$$\text{sujeto a: } y_j(a^T x_j + b) \geq 1$$

a) Pruebe que la dirección óptima está dada por $a^* = \sum_{j=1}^m \mu_j^* y_j x_j$ donde $\mu^* := (\mu_1, \ldots, \mu_m^*)^T$, es solución de (4) con $\mathbf{1} = (1, \ldots, 1)^T \in \mathbb{R}^m$, $y = (y_1, \ldots, y_m)^T \in \mathbb{R}^m$, y $Q \in \mathbb{R}^{m \times m}$ a determinar

$$\min_{\mu \geq 0} \mu^T Q \mu - \mathbf{1}^T \mu \qquad (4)$$
$$\text{sujeto a: } y^T \mu = 0$$

b) Justifique la siguiente propiedad: La dirección óptima a^* es una combinación lineal de "supporting vectors" x_j que están sobre el borde de la banda.

El planteo del problema anterior asume que la muestra es separable. Si esto no se cumple entonces pueden agregarse variables extra ϵ_j para permitir que algunos puntos del conjunto de entrenamiento queden mal clasificados

$$(a^*, b^*, \epsilon^*) = \arg \min_{(a,b,\epsilon)} \frac{1}{2}\|a\|^2 + c \sum_{j=1}^m \epsilon_j \qquad (5)$$
$$\text{sujeto a: } y_j(a^T x_j + b) \geq 1 - \epsilon_j$$
$$\epsilon_j \geq 0$$

c) (opcional) Pruebe que en este caso los multiplicadores que determinan la dirección de la recta son solución de (nótese que en este caso se agrega la restricción $\mu_j \leq c, \ j = 1, \ldots, m$)

$$\min_{\mu \in [0,c]^m} \mu^T Q \mu - \mathbf{1}^T \mu \qquad (6)$$
$$\text{sujeto a: } y^T \mu = 0$$

d) (opcional) Muestre que los x_j bien clasificados que caen por fuera de la banda corresponden a $\mu_j = 0$.

e) (opcional) Muestre que los x_j mal clasificados o los que caen dentro de la banda corresponden a $\mu_j = c$

f) (opcional) Resuelva (5) para los datos adjuntos x.asc y y.asc, grafique la solución, y verifique que se cumplen las propiedades de las partes b) y c).

Ejercicio 4 - Gestión óptima de la demanda de energía eléctrica

Considere el ejemplo de una red de tres barras de la figura 1 con dos generadores y dos cargas. Los costos de generación en las barras 1 y 2 son los que muestran en la figura 1 y están dados por

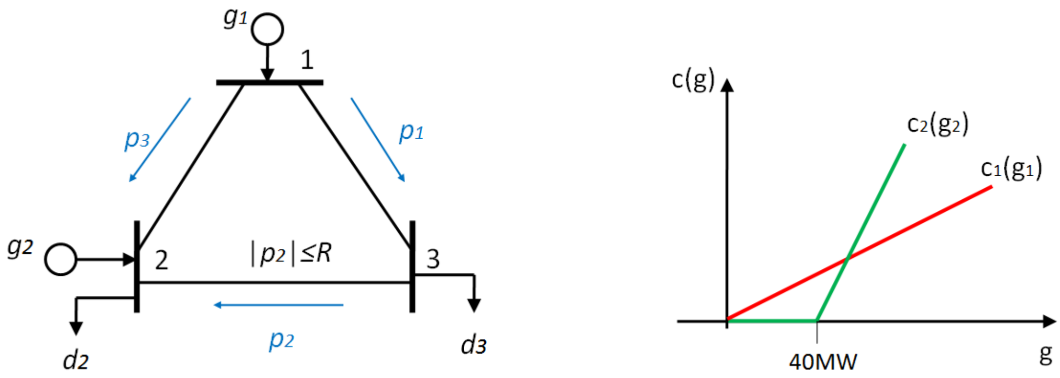


Figura 1: Red eléctrica de tres barras y costos asociados de generación.

$c_1(g_1) = g_1$ y $c_2(g2) = \max\{0, 4(g_2 - 40MW)\}$

El problema de optimización que modela esta red es (ver notas del curso)

$$\min_{(p_1,p_2,p_3,g_1\geq 0,g_2\geq 0)} c_1(g_1) + c_2(g_2) \tag{7}$$

$$\text{sujeto a: } p_1 + p_3 = g_1 \tag{8}$$

$$g_2 + p_3 + p_2 = d_2 \tag{9}$$

$$p_1 - p_2 = d_3 \tag{10}$$

$$p_3 - p_1 - p_2 = 0 \tag{11}$$

$$|p_2| \leq R \tag{12}$$

donde $R = 30MW$, $\bar{d}_3 = 10MW$ y d_2 se deja como parámetro del problema.

a) Reformule (7) como un programa lineal (LP) introduciendo la variable t (slack) tal que $c_2(g_2) \leq t$.

b) Resuelva numéricamente (LP) para cada valor de $d_2 \in \{0 : 1 : 200\}$ y grafique g_1, g_2, p_2 y λ en función de d_2, siendo λ el multiplicador asociado a (10). Se sugiere usar Julia/Convex-Matlab/CVX-Python/cvxpy.

c) Interprete las gráficas de la parte anterior, explicando los valores de los multiplicadores y justificando los puntos de inflexión. Determine para que valores de d_2 es conveniente ofrecer energía a precio nulo en la barra 3.

Ejercicio 5 - *Relajación de problemas binarios*

Considérense los siguientes problemas con $Q > 0$.

$$(PB) \min_{x \in \mathbb{R}^n} x'Px + 2q'x + r, \qquad (PC) \min_{x \in \mathbb{R}^n} x'Px + 2q'x + r$$
$$s.to : x_i^2 = 1, \ i = 1 \ldots, n \qquad s.to : x_i^2 \leq 1, \ i = 1 \ldots, n$$

a) Escriba el problema dual (DB) correspondiente a (PB) como un problema SDP. Para esta parte se puede usar sin demostrar el resultado en las notas del curso para problemas cuadráticos no convexos.

b) Considérese la solución subóptima del problema binario (PB) que se obtiene aproximando la solución de (PC) a las variables binarias más cercanas ($x_i = \text{sign}(x_i)$). Llamaremos (PA) a este método.

c) Ordene el mínimo valor funcional de (PB), el máximo de (DB), y el obtenido mediante (PA). Nótese que de este orden puede obtenerse un intervalo en el cual debe caer el valor funcional óptimo de (PB) y una cota superior para el gap por utilizar la solución subóptima (PA).

De aquí en más se pasará a resolver los problemas numéricamente usando la matriz $P \in \mathbb{R}^{10\times10}$, vector $q \in \mathbb{R}^{10\times1}$ y escalar $r \in \mathbb{R}$ que se encuentran en los archivos adjuntos *P.asc*, *q.asc* y *r.asc*. Para la resolución de estos problemas puede utilizarse Matlab/CVX, Julia/Convex, o Python/cvxpy.

d) Resuelva (PA), y (DB) verificando que se respeta el orden obtenido en la parte c).

e) Resuelva (PB) probando todas las $1024 = 2^{10}$ combinaciones posibles de valores binarios y verifique que respeta el orden de la parte c=.

f) (opcional) Plantee (DC) y mediante comparación con (DB) intercale los mínimos y máximos valores funcionales de (PC) y (DC) en el orden de la parte c)

g) (opcional) Resuelva (DC) y verifique que se cumple dualidad fuerte con (PC).

Ejercicio 6 - Óptimo como punto silla del Lagrangeano (Opcional)

Sea una función de dos variables $\varphi(x,z) : X \times Z \to \mathbb{R}$.

a) Pruebe la desigualdad $\sup_x \inf_z \varphi(x,z) \leq \inf_x \sup_z \varphi(x,z)$ [1,p.151].

b) Suponga que la función $\varphi(x,z)$ es diferenciable, convexa en $x \in X$ y cóncava $z \in Z$, con X y Z conjuntos convexos. Demuestre que un punto $(\tilde{x},\tilde{z})$ tal que $\nabla\varphi(\tilde{x},\tilde{z}) = 0$ es punto silla, esto es,

$$\varphi(\tilde{x},z) \leq \varphi(\tilde{x},\tilde{z}) \leq \varphi(x,\tilde{z}) \tag{13}$$

c) Pruebe que (13) implica

$$\sup_x \inf_z \varphi(x,z) \geq \inf_x \sup_z \varphi(x,z), \implies \sup_x \inf_z \varphi(x,z) = \inf_x \sup_z \varphi(x,z), \tag{14}$$

con lo que supremo e ínfimo pueden intercambiarse sin afectar el resultado.

Sea $\varphi(x,z)$ el Lagrangeano del problema (15) con f convexa y diferenciable respecto a x, y sean $z \in \mathbb{R}^m$ los multiplicadores asociados a (16)

$$\min_{x\in X} f(x) \tag{15}$$

$$\text{sujeto a: } Ax - b = 0 \in \mathbb{R}^m \tag{16}$$

d) Observe que el lado derecho de la igualdad en (14) es el máximo del dual de (15) y muestre que el lado derecho corresponde al mínimo del primario (ver [1,p.237]).

e) Muestre que se cumplen las condiciones en b). Luego (14) garantiza la dualidad fuerte en (15).

Ejercicio 7 - Dualidad en problemas de transporte (Opcional [2,p.440])

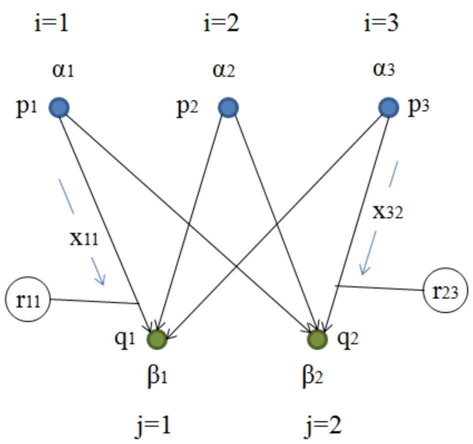


Figura 2: Esquema de transporte desde $n = 3$ puntos de oferta a $m = 2$ puntos de demanda

Se desea transportar un bien a mínimo costo desde n puntos de oferta a m puntos de demanda según se observa esquemáticamente en la figura 2. El bien tiene un precio p_i en el sitio de oferta $i = 1, \dots, n$ y un precio q_j en el sitio de demanda $j = 1, \dots, m$. Cada transportista incurre en un costo $r_{ij}x_{ij}$ por llevar una cantidad x_{ij} desde i a j con costo marginal constante r_{ij}, con lo que maximiza su excedente trasladando la cantidad

$$x_{ij}(p_i, q_j) = \begin{cases} 0, \text{ si } r_{ij} > q_j - p_i \\ \infty, \text{ si } r_{ij} < q_j - p_i \\ x \geq 0, \text{ si } r_{ij} = q_j - p_i \end{cases} \tag{17}$$

con x_{ij} indeterminado cuando el costo marginal iguala a la diferencia de precios de compra y venta.

Por otra parte se plantea el problema de minimizar los costos globales de transporte

$$(PT) \min_{x_{i,j} \geq 0} \sum_{i=1}^{n} \sum_{j=1}^{m} r_{ij}x_{ij} \tag{18}$$

$$\text{sujeto a: } \sum_{j=1}^{m} x_{ij} \leq \alpha_i, \ i = 1, \ldots, n \tag{19}$$

$$\sum_{i=1}^{n} x_{ij} \geq \beta_j, \ j = 1, \ldots, m \tag{20}$$

donde α_i y β_j en (19) y (20) son dados y representan la oferta y demanda en los sitios i y j, respectivamente. Para que el problema sea factible se asume con $\sum_{j=1}^{m} \beta_j \leq \sum_{i=1}^{n} \alpha_i$.

a) Plantee el Lagrangeano dualizando (19) y (20) y dejando $X = \{x_{ij} \geq 0\}$ como conjunto factible.

b) Muestre que si se igualan los multiplicadores a los precios p_i y q_j la estrategia (17) minimiza el Lagrangeano en X.

c) Obtenga la función dual y plantee el problema dual (DT) correspondiente a (PT) en (18)-(20).

d) Argumente dualidad fuerte y deduzca de ello que los precios óptimos son tales que $q_j^* - p_i^* = r_{ij}$ si $x_{ij}^* > 0$, descartando el segundo caso en (17).

e) Para los costos marginales $r_{11} = 1$, $r_{12} = 2$, $r_{21} = 5$, $r_{22} = 3$, $r_{31} = 3$, y $r_{22} = 2$, las ofertas $\alpha = [3, 4, 5]$, y las demandas $\beta = [6, 6]$, implemente (18) en Matlab o Julia, verificando que la estrategia (17) es óptima y que se cumple la condición obtenida en la parte anterior.

Referencias

[1] S. Boyd, V. Vanderberghe *Convex Optimization,* Cambridge University Press, 2006.

[2] D. P. Bertsekas, *Nonlinear programming, 3rd ed.,* Athena Scientific, 2017.