ICS 2121 Metodos de Optimización, Sem 2018-2
Prof. Jorge Vera

Interrogación 2. Solución.

Pregunta 1 (14 puntos):

a) (7 pts) Considere el poliedro definido por el siguiente sistema de desigualdades:

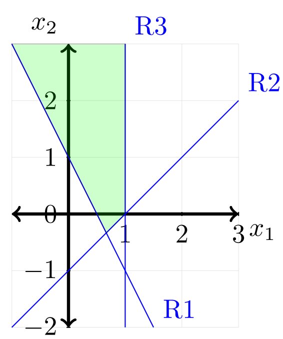
$$\begin{array}{lll} 2x_1 & +x_2 & \geq 1 \\ x_1 & -x_2 & \leq 1 \\ & x_1 & \leq 1 \end{array}$$

i) (3 pts) Usando solamente el concepto de cono de recesion, muestre que este poliedro no es acotado.
Respuesta: El cono de recesion asociado al poliedro viene dado por el siguiente conjunto de desigualdades:

$$x_2 \geq -2x_1$$
$$x_2 \geq x_1$$
$$x_1 \leq 0$$

Basta con encontrar cualquier punto distinto del vector cero que cumpla con el conjunto de desigualdades. Como el $(0,1)$ cumple con todas las desigualdades, el poliedro es no acotado.

ii) (4 pts) El punto $(0,2)$ es evidentemente factible, como se deduce por inspeccion. Escriba explicitamente ese punto usando la descomposicion en puntos y rayos extremos para el poliedro de mas arriba.
Respuesta:



Para escribir el punto $(0,2)$ utilizando el teorema de Minkowski, primero debemos encontrar los vertices y rayos que definen el poliedro. Estos son:

- Vertices: $A = (\frac{2}{3}, -\frac{1}{3})$, $\quad B = (1,0)$
- Rayos: $y^1 = (0,1)$, $\quad y^2 = (-1,2)$

Ahora, debemos encontrar los valores de $\lambda_1, \lambda_2, \mu_1$ y μ_2 tales que:

$$\begin{pmatrix} 0 \\ 2 \end{pmatrix} = \lambda_1 \begin{pmatrix} 2/3 \\ -1/3 \end{pmatrix} + \lambda_2 \begin{pmatrix} 1 \\ 0 \end{pmatrix} + \mu_1 \begin{pmatrix} 0 \\ 1 \end{pmatrix} + \mu_2 \begin{pmatrix} -1 \\ 2 \end{pmatrix}$$
$$\lambda_1 + \lambda_2 = 1$$
$$\mu_1, \mu_2 \geq 0$$

b) (7 pts) Considere el siguiente problema de optimizacion no lineal:

$$P)\quad \begin{aligned} z^* = \max \quad & a^T x \\ s.a. \quad & x^T x \leq 1 \end{aligned}$$

donde $a \in \mathbb{R}^n$, $a \neq 0$.

i) (4 pts) Determine y escriba el problema dual a P). (Indicacion: tenga cuidado con el hecho que el problema esta escrito como maximizacion)

Respuesta:

$$P)\quad \begin{aligned} z^* = \min \quad & -a^T x \\ s.a. \quad & x^T x \leq 1 \end{aligned}$$

El lagrangeano y funcion dual asociadas al problema son:

$$\begin{aligned} L(x, \lambda) &= -a^T x + \lambda(x^T x - 1) \\ \implies \theta(\lambda) &= \min_{x \in \mathbb{R}} \{-a^T x + \lambda(x^T x - 1)\} \end{aligned}$$

Ahora, encontraremos el valor de x que minimiza la expresion, derivando e igualando a 0:

$$\begin{aligned} \frac{\partial L(x, \lambda)}{\partial x} &= 0 \implies x(\lambda) = \frac{a}{2\lambda} \\ \implies \theta(\lambda) &= -a^T x(\lambda) + \lambda(x(\lambda)^T x(\lambda) - 1) \\ &= \frac{a^T a}{2\lambda} + \lambda \frac{a^T a}{4\lambda^2} - \lambda \\ &= -\frac{a^T a}{4\lambda} - \lambda \end{aligned}$$

Luego, el problema dual viene dada por la maximizacion en las variables λ de la funcion dual, esto es:

$$\begin{aligned} \max_{\lambda \geq 0} \theta(\lambda) &= \max_{\lambda \geq 0} \{-\frac{a^T a}{4\lambda} - \lambda\} \\ &= \min_{\lambda \geq 0} \{\frac{a^T a}{4\lambda} + \lambda\} \end{aligned}$$

ii) (3 pts) Â¿Es el valor optimo de P) igual al valor optimo del problema dual? Justifique con precision su respuesta.

Respuesta: Si, ya que se cumplen las condiciones del Teorema Fuerte de Dualidad:

- Problema convexo: Funcion objetivo lineal y $x^T x$ es convexo x.
- Condicion de Slater: Por ejemplo, $x = 0$ cumple con $x^T x < 1$.

iii) (4 pts de bono) Usando el problema dual desarrollado en ii), muestre que $z^* = \|a\|_2$.

Respuesta: Tenemos que encontrar el valor de λ que minimiza la expresion $f(\lambda) = \frac{a^T a}{4\lambda} + \lambda$. Derivando e igualando a 0:

$$\begin{aligned} \frac{\partial f(\lambda)}{\partial \lambda} &= 0 \leftrightarrow -\frac{a^T a}{4\lambda^2} + 1 = 0 \\ \implies \lambda &= \frac{\sqrt{a^T a}}{2} \end{aligned}$$

Reemplazando en la funcion objetivo obtenemos:

$$\begin{aligned} z* &= \frac{a^T a}{2\sqrt{a^T a}} + \frac{\sqrt{a^T a}}{2} \\ &= \frac{a^T a}{\sqrt{a^T a}} \\ &= \sqrt{a^T a} \\ &= \|a\|_2 \end{aligned}$$

Pregunta 2 (14 puntos):
Considere el Metodo de Descomposicion de Dantzig-Wolfe aplicado a un problema de Programacion Lineal con la estructura que presentamos en clases:

$$
\begin{array}{llcll}
\min & c_1^T x^1 & + \cdots + & c_K^T x^K & \\
s.t. & A_1 x^1 & + \cdots + & A_K x^K & = b \\
 & D_1 x^1 & & & = d_1 \\
 & \vdots & \ddots & \vdots & \vdots \\
 & & & D_K x^K & = d_K \\
 & x_1 \geq 0 & & x_K \geq 0 &
\end{array}
$$

donde $A_k \in \mathbb{R}^{m \times n_k}$, $D_k \in \mathbb{R}^{p_k \times n_k}$, $d_k \in \mathbb{R}^{p_k}$, $c_k \in \mathbb{R}^{n_k}$, $k = 1, \ldots, K$, $b \in \mathbb{R}^m$.
Como usted sabe, la descomposiciï¿½n parte considerando los poliedros $P_k = \{x \in \mathbb{R}^{n_k} : D_k x = d_k, x \geq 0\}$. En el desarrollo en clases asumimos que estos conjuntos eran acotados. Suponga ahora que son poliedros generales, es decir, pueden haber rayos en los correspondientes conos de recesiï¿½n.

a) (8 pts) Escriba el correspondiente problema transformado que se debe usar para la Descomposicion de Dantzig-Wolfe (el problema que resulta de reemplazar los vertices y rayos de los poliedros P_k en la formulacion original).
Respuesta: Para cada P_k tenemos vertices $\{v_k^i\}$ y rayos $\{y_k^i\}$.

$$\forall k = 1, ..., K \quad x_k = \sum_{i=1}^{n_k} v_k^i \lambda^i + \sum_{j=1}^{m_k} y_k^j \mu_k^j$$

$$\sum_{i=1}^{n_k} \lambda_k^i = 1 \quad \mu_k^j, \lambda_k^i \geq 0$$

Al reemplazar, quedan los siguientes costos:

$$c^T x_k = c^T \left(\sum_{i=1}^{n_k} v_k^i \lambda^i + \sum_{j=1}^{m_k} y_k^j \mu_k^j\right) = \sum_{i=1}^{n_k} \lambda^i c_k^T v_k^i + \sum_{j=1}^{m_k} \mu_k^j c_k^T y_k^j$$

Al aplicar el mismo reemplazo en las restricciones complicantes, el problema queda:

$$\min \quad \sum_{i=1}^{n_k} \lambda^i c_k^T v_k^i + \sum_{j=1}^{m_k} \mu_k^j c_k^T y_k^j$$

$$s.a. \quad \sum_{i=1}^{n_k} \lambda^i A_k v_k^i + \sum_{j=1}^{m_k} \mu_k^j A_k y_k^j$$

$$\sum_{i=1}^{n_k} \lambda_i^k = 1 \quad \forall k = 1, ..., K$$

$$\lambda_i^k \geq 0 \quad \forall k = 1, ..., K \quad \forall i = 1, ..., n_k$$

$$\mu_j^k \geq 0 \quad \forall k = 1, ..., K \quad \forall j = 1, ..., m_k$$

Cabe destacar que, dado que usamos la informacion de cada bloque para formular los x_k como lo estamos haciendo, no es necesario expresar dichas restricciones, ya que la igualdad siempre se cumple.

b) (6 pts) Escriba el problema de "Pricing" que se debe resolver cuando se aborda el problema transformado mediante Generacion de Columnas, y explique como ese problema determina una columna para ser agregada al problema maestro.
Respuesta: El problema de "pricing" requiere calcular los costos reducidos, por lo que notamos π como las variables duales obtenidas del maestro para las restricciones complicantes y t las asociadas a

la suma de los λ.
Para el vertice i correspondiente al poliedro k, se tiene lo siguiente:

$$\bar{c_{ki}} = c_{ki} - \pi^T A_{ki} = c_k^T v_{ki} - \pi^T A_k^T v_{ki} - t_k$$

Para el rayo j correspondiente al poliedro k, se tiene tambien:

$$\bar{c_{kj}} = c_k^T y_{kj} - \pi^T A_k^T y_{kj} - t_k$$

Luego, se busca un vertice **o rayo** que entregue costos reducidos negativos. Para esto, se formula el problema satelite:

$$\eta_k = \min(c_{ki} - \pi^T A_k^T \pi)^T u$$
$$s.a. D_k^T u = d_k$$
$$u \geq 0$$

Si el problema retorna un optimo tal que $\eta_k - t_k \leq 0$, este corresponde a un vertice y se continua iterando entre maestro y satelite. Sin embargo, si el problema es no acotado, se encuentra un rayo que no permite agregar una columna.

Pregunta 3 (12 puntos)
Considere un problema de corte de piezas similar al visto en clases para explicar Generacion de Columnas: disponemos de listones de madera de largo L y tenemos que cortar de ellos una cantidad de piezas de p distintos tipos de diferentes largos. Sean $l_1, \ldots, l_p$ los largos. Para cada tipo de pieza hay demandas que tenemos que satisfacer. Sean $d_1, \ldots, d_p$ las demandas por las distintas piezas.
Siguiendo el mismo desarrollo que estudiamos en el curso para el problema de corte de piezas, supongamos que tenemos identificadas distintas formas de cortar las piezas, esos seran los "patrones de corte". El patron j genera a_{ij} piezas de largo l_i. Pero ahora, adicionalmente tendremos que cada patron tiene un costo asociado, que se relaciona al proceso de corte. Especificamente, existe un costo asociado al manejo del proceso de corte y que es proporcional al numero de piezas cortadas de un liston. El costo de corte del patron es igual a $k \times \gamma$, donde γ es un costo unitario por pieza generada y k es el numero de piezas que se obtienen. Por ejemplo, si $L = 100$ y tenemos 5 tipos de piezas con largos $(30, 5, 10, 20, 40)$, y un patron es $(0, 2, 6, 0, 0)$, esto significa que se obtienen 8 piezas (2 de largo 5 y 6 de largo 10) y el costo es $8 \times \gamma$. De esta manera, cada patron genera un costo, que denotaremos α_j, asociado a lo anterior. Adicionalmente, cada liston tiene un costo unitario de compra igual a β.
El problema, es, entonces, determinar la cantidad de listones que seran cortados con cada patron de corte de modo que el costo total incurrido (costo de compra de los listones y costo del patron) sea minimo. Para esto, se puede usar el siguiente modelo de Optimizacion donde la variable principal es x_j, igual al numero de listones a cortar con el patron j, y se asume que existen r patrones de corte:

$$\begin{array}{ll} \min & \sum_{j=1}^{r} (\alpha_j + \beta) x_j \\ s.a. & \sum_{j=1}^{r} a_{ij} x_j \geq d_i \; i = 1, \ldots, p \\ & x_j \geq 0, \text{ entero} \end{array}$$

Desarrolle el detalle de como resolver la relajacion lineal del modelo anterior mediante Generacion de Columnas, ya que no es posible conocer simultaneamente todos los patrones de corte. Sea explicito, entre otras, en los problemas "Maestro" y "Satelite".

Respuesta:
La relajacion lineal a resolver es:

$$\begin{array}{ll} \min & \sum_{j=1}^{r} (\alpha_j + \beta) x_j \\ s.a. & \sum_{j=1}^{r} a_{ij} x_j \geq d_i \; i = 1, \ldots, p \\ & x_j \geq 0, \text{ entero} \end{array}$$

Sea $C_j = \alpha_j + \beta$ el patron que entra a la base y determina el precio, entonces:

$$\overline{C_j} = C_j - \pi^T A_j = \alpha_j + \beta - \pi^T A_j = (\alpha_j - \pi^T A_j) + \beta$$

Luego, hay que buscar un patron j tal que se cumpla $\overline{C_j} < 0$.Se puede reescribir $\overline{C_j} < 0$ de tal forma que:

$$\overline{C_j} = (\alpha_j - \sum_{i=1}^{p} \pi^T a_{ij}) + \beta$$

Se debe ener en cuanta que $\sum_{i=1}^{p} l_i a_{ij} \leq L$. Finalmente el problema satelite queda de la forma:

$$\begin{aligned} \eta = \min \quad & \sum_{i=1}^{p} u_i - \sum_{i=1}^{p} \pi_i u_i \\ s.a. \quad & \sum_{i=1}^{p} l_i u_i \leq L \\ & u_i \geq 0, \text{ entero} \end{aligned}$$

Este problema de la mochila calcula un patron valido de $\eta + \beta < 0$. El maestro es:

$$\begin{aligned} \min \quad & \sum_{j \in J} \alpha_j x_j \\ s.a. \quad & \sum_{j \in J} a_{ij} x_j \geq d_i \ \forall i \\ & x_j \geq 0 \end{aligned}$$

Donde J es el conjutno de columnas hasta el momento.

Pregunta 4 (14 puntos):

Una empresa del area agricola posee varios campos en los cuales hay disponible cantidades de un cultivo que debe ser cosechado y transportado hasta ciertas plantas de procesamiento, donde hay una demanda conocida. El problema es que la empresa debe construir caminos para poder llevar adelante toda la operacion de transporte, ya que de un año al siguiente, esos caminos son destruidos por las condiciones climaticas. La empresa ha decidido representar su situacion en un grafo (N, A), donde N es el conjunto de los nodos (que pueden ser campos, plantas de procesamiento o, simplemente, puntos de paso), y A son los posibles caminos que podrian construirse. Vamos a asumir que en cada nodo hay una oferta (o demanda, segun el signo) igual a $d_i, i \in N$ y supondremos que la oferta esta balanceada con la demanda. La empresa ha formulado un modelo de Optimizacion usando variables x_{ij}, que representan la cantidad transportada entre el nodo i y el nodo j, y variables binarias y_{ij} que indican si el camino (i, j) es construido o no. Obviamente, si un camino no es construido, no puede haber transporte por ahi. Sea c_{ij} el costo unitario de transporte por el camino (i, j), y K_{ij} el costo fijo de inversion por construir el camino (i, j). Existe tambien una capacidad U_{ij} de transporte en el camino (i, j), en caso que este sea construido. Adicionalmente, por razones ambientales, hay restricciones a los caminos que pueden ser construidos. Mas especificamente, existen algunas areas geograficas en donde el numero de caminos construidos no puede exceder un limite dado. Esto es representado por p conjuntos $B_1, \ldots, B_p$, donde $B_k \subset A, k = 1, \ldots, p$. La restriccion es que para el area representada por los posibles caminos del conjunto B_k, no pueden construirse mas de n_k caminos. El modelo es el siguiente:

$$
\begin{array}{llr}
\min & \sum_{(i,j)\in A} c_{ij}x_{ij} + \sum_{(i,j)\in A} K_{ij}y_{ij} & \\
s.a. & \sum_{(i,j)\in A} x_{ij} - \sum_{(k,i)\in A} x_{ki} = d_i, \quad i \in N & (1)\\
& x_{ij} \le U_{ij}y_{ij}, \quad (i,j)\in A & (2)\\
& \sum_{(i,j)\in B_k} y_{ij} \le n_k \quad k = 1,\dots,p & (3)\\
& x_{ij} \ge 0, y_{i,j} \in \{0,1\} \quad (i,j)\in A & (4)
\end{array}
$$

a) (3 pts) Identifique estructuras de variables complicantes en este problema, que hagan suceptible que el problema sea resuelto mediante Descomposicion de Benders. Muestre esta estructura en un esquema en el que indentifique los bloques adecuados para la descomposicion.

Respuesta: Las y_{ij} son las variables complicantes, por lo que podemos definir la estructura de Benders en este problema de la siguiente forma

$$
\begin{array}{lcccl}
\min & \mathbf{Y} & + & \mathbf{X} & \\
s.a. & \mathbf{A} & & & \le \mathbf{b}\\
& \mathbf{E^1} & + & \mathbf{D^1} & \le \mathbf{e^1}\\
& \mathbf{E^2} & + & \mathbf{D^2} & = \mathbf{e^2}
\end{array}
$$

Con $\mathbf{Y} = \sum_{(i,j)\in A} K_{ij}y_{ij}$, $\mathbf{X} = \sum_{(i,j)\in A} c_{ij}x_{ij}$, $\mathbf{A} = \sum_{(i,j)\in B_k} y_{ij} \forall k = 1,\dots,p$ y $\mathbf{b} = n_k \forall k = 1,\dots,p$. Finalmente, las matrices E, D y el vector e se definen por

$$
E = \begin{bmatrix} E^1 \\ - \\ E^2 \end{bmatrix} = \begin{bmatrix} -U_{ij} & & \\ & \ddots & \\ & & -U_{i'j'} \\ -- & -- & -- \\ & 0 & \end{bmatrix}, D = \begin{bmatrix} D^1 \\ - \\ D^2 \end{bmatrix} = \begin{bmatrix} I \\ ---------- \\ \sum_{(i,j)\in A} x_{ij} - \sum_{(k,i)\in A} x_{ki} \end{bmatrix}, e = \begin{bmatrix} e^1 \\ - \\ e^2 \end{bmatrix} = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ - \\ d_i \end{bmatrix}
$$

b) (8 pts) Escriba explicitamente los problemas maestros y satelites para la Descomposicion de Benders e indique como funciona el algoritmo (Ind: le puede ser util usar el equema de la parte a), usando nombres simbolicos para esos bloques, pero al final debe expresar todo en terminos de los datos originales.)

Respuesta: Cuando se hace Benders, el problema interior es

$$
\begin{array}{llr}
\min & \sum_{(i,j)\in A} c_{ij}x_{ij} & \\
s.a. & \sum_{(i,j)\in A} x_{ij} - \sum_{(k,i)\in A} x_{ki} = d_i, \quad i \in N & (1)\\
& x_{ij} \le U_{ij}y_{ij}, \quad (i,j)\in A & (2)\\
& x_{ij} \ge 0 \quad (i,j)\in A & (3)
\end{array}
$$

Este problema es de flujo de costo minimio capacitado y su dual es

$$
\begin{array}{ll}
\max & \sum_{(i\in N} \pi_i d_i + \sum_{(i,j)\in A} W_{ij}U_{ij}y_{ij}\\
s.a. & \pi_i - \pi_j + W_{ij} \le C_{ij} \quad (i,j)\in A\\
(PS) & W_{ij} \ge 0 \quad (i,j)\in A
\end{array}
$$

Con π_i y W_{ij} como variables duales. Este problema es la base del satelite.

Por su parte, el problema maestro es:

$$
\begin{array}{lrl}
\min & \sum_{(i,j)\in A} K_{ij}y_{ij}+\gamma & \\
s.a. & \sum_{(i,j)\in B_k} y_{ij}\leq n_k & k=1,\ldots,p\\
& \sum_{(i,j)\in A} W^l_{ij}U_{ij}y_{ij}\leq \gamma-\sum_{(i\in N} \pi^l_i d_i & l=1,\ldots,q\\
(PM) & x_{ij}\geq 0, y_{i,j}\in\{0,1\} & (i,j)\in A
\end{array}
$$

Donde (W^l,π^l) corresponde al l-esimo punto extremo de (PS)pasando por la iteracion en el satelite. En el satelite (PS), y_{ij} llega fijado como solucion del maestro (PM)

c) (3 pts) Un amigo le dice que hay tambien en el problema una estructura de restricciones complicantes inducida por las restricciones (2) y que podria usar descomposicion de Dantzig-Wolfe. Â¿Esta usted de acuerdo y usaria Dantzig-Wolfe para este problema? Explique con claridad su respuesta.

Respuesta: Efectivamente es verdad que las restricciones (2) establecen una relacion entre las variables x y las y. Por otro lado, las restricciones (1) solo involucran a las variables x, mientras que las restricciones (3) solo involucran a las y. Eso significa que, si se ve desde ese punto de vista, el problema tiene dos bloques y restricciones comunes, lo que lo hace tambien suceptible de ser abordado mediante Dantzig-Wolfe.

Respecto a si realmente debeira usarse DW aquí, formalmente no hay problemas. Si podría ser objetable el que el problema se separa sólo en dos bloques. La idea de la descomposición es que existan varios bloques que simultáneamente aporten columnas.

Indicacion: Note que el problema no esta exactamente en el mismo formato que estudiamos en clases (tiene mezcla de = y ≤), pero eso no importa. Dada la diferencia de formato, le puede ser util el siguiente recordatorio para duales de Programacion Lineal:

Minimizacion	Maximizacion
Variables	Restricciones
≥ 0	$\leq$
≤ 0	$\geq$
Irrestrictas	$=$
Restricciones	Variables
$\leq$	≤ 0
$\geq$	≥ 0
$=$	Irrestrictas