ICS 2121 Métodos de Optimización, Sem 2015-2
Prof. Jorge Vera

Interrogación 3. SOLUCION

Pregunta 1 (12 puntos):

a) (3 pts) Consideremos un problema no lineal general de la forma

$$P) \quad \begin{array}{rl} z^* = \min & f(x) \\ s.a. & g_i(x) \leq b_i, i = 1, \ldots, m \\ & x \in C \end{array}$$

donde f y las funciones g_i son convexas y C es un conjunto convexo. Entonces el valor óptimo del dual de este problema es siempre igual al valor z^* del P). ¿Es verdad esta afirmación o no? Justifique con claridad.

Respuesta: En el caso general convexo se requiere además la condición de Slater.

b) (3 pts) Supongamos que hemos formulado el dual del problema P) de la parte a) y queremos resolverlo maximizando la función dual, $\theta(\lambda)$, mediate el método del "supgradiente" ¿Por qué no se quiere hacer "linesearch" cuando se aplica ese método?

Respuesta: Cualquier método de linesearch requeire evaluar la función en cuestión, $\theta(\lambda)$, en este caso, repetidas veces. En este caso evaluar $\theta(\lambda)$ implica resolver un problema de optimización, lo cual puede ser difícil o largo. Por esa razón no queremos iterar en un linsearch.

c) (3 pts) Explique qué significa hacer "multiple pricing" en Generación de Columnas.

Respuesta: Es agregar varias columnas de costo reducido negativo a la vez y trabajar en el Mestro con ellas hasta que se agoten.

d) (3 pts) Considere nuevamente un problema general como el P) de la parte a). Explique qué dice el Teorema Débil de Dualidad.

Respuesta. Dice que el valor del dual acota por abajo al valor del primal.

Pregunta 2 (12 puntos)

Una empresa necesita definir la cantidad de empleados que debe asignar en distintos turnos de trabajo para un periodo de 24 horas. Los turnos están diseñados para que sea posible atender ciertas ventanas de tiempo en las que se requieren cantidades mínimas de personal. Por ejemplo, la siguiente tabla muestra un esquema de cuatro turnos y 8 ventanas de tiempo:

TABLE 2.1 Time Windows for Shift Workers

	Shift				
Time Window	1	2	3	4	Workers Required
6 a.m.–9 a.m.	X			X	55
9 a.m.–12 noon	X				46
12 noon–3 p.m.	X	X			59
3 p.m.–6 p.m.		X			23
6 p.m.–9 p.m.		X	X		60
9 p.m.–12 a.m.			X		38
12 a.m.–3 a.m.			X	X	20
3 a.m.–6 a.m.				X	30
Wage rate per 9 h shift	\$135	\$140	\$190	\$188	

Cada turno consiste en trabajar en algunas de las ventanas de tiempo, con algunas condiciones: el turno debe estar formado por 3 ventanas de tiempo (no necesariamente consecutivas) y para las tres últimas ventanas, las de la noche y madrugada, de 9 p.m. a 6 a.m., pueden usarse a lo más dos de esas (por ejemplo, el turno 2 cubre tres ventanas de tiempo y corresponde a trabajar entre el mediodía y 9 p.m. mientras que el turno 4, como ya cubrió la noche sólo puede tener otra ventana separada más temprano). En un caso general, supongamos que hay T ventanas de tiempo y n turnos. Sea d_t, la cantidad mínima de empleados requerida para la ventana de tiempo t, $t = 1, \ldots, T$. El salario por trabajar en la ventana t es s_t y el salario total para un empleado que trabaja en un turno j, y que denotamos w_j, es la suma de los salarios s_t para las ventanas que forman el turno. La estructura de los turnos (como en la tabla, por ejemplo) la representaremos por un *parámetro* a_{jt} el cual tendrá valor 1 si la ventana de tiempo t, forma parte del turno j, y 0 si no.
La situación se modela usando una variable x_j, entera, que corresponde a la cantidad de personas que se deben contratar para el turno j. El modelo que determina la cantidad de personas a contratar por cada turno, de modo que el costo sea mínimo y cubra las necesidades es:

$$\begin{array}{ll} \min & \sum_{j=1}^{n} w_j x_j \\ s.a. & \sum_{j=1}^{n} a_{jt} x_j \geq d_t, t = 1, \ldots, T. \\ & x_j \geq 0 \text{ entero }, j = 1, \ldots, n. \end{array}$$

a) (4 pts) En primer lugar, argumente que las reglas descritas que deben cumplir los turnos, permiten efectivamente generar suficientes turnos para cubrir todas las ventanas. Escriba explícitamente en forma matemática las condiciones que deben cumplir los parámetros a_{jt}, para que el turno j sea válido.

Respuesta: El mismo ejemplo de la tabla muestra que hay suficientes patrones para cubrir los requerimientos.

Las condiciones que deben cumplir las asignacíones de turnos, que están dadas por los parámetros a_{jt}, son:

$$\sum_{t=1}^{8} a_{jt} = 3,\ a_{j6} + a_{j7} + a_{j8} \leq 2,\ a_{jt} \in \{0, 1\}$$

b) (8 pts) Ahora explique cómo se puede resolver la relajación lineal del problema (es decir, sin incluir la restricción de enteros, sólo $x_j \geq 0$) usando Generación de Columnas. En particular, describa con precisión el problema Maestro y el Satélite que se deben usar.

Respuesta: Aquí tienen que dar el detalle del método. La parte menos trivial es lo del satélite ya que si π son los multiplicadores del Simplex de una iteración, tenemos que busca una columna con costo reducido negativo. El costo reducido de una columna es:

$$\bar{w}_j = w_j - \pi^T A_j = \sum_{t=1}^{8} a_{jt} s_t - \sum_{t=1}^{8} a_{jt} \pi_t = \sum_{t=1}^{8} a_{jt}(s_t - \pi_t)$$

Luego, el satélite es de la forma:

$$\begin{array}{ll} \min & \sum_{t=1}^{8} u_t(s_t - \pi_t) \\ s.a. & \sum_{t=1}^{8} u_t = 3 \\ & u_6 + u_7 + u_8 \leq 2 \\ & u_j \in \{0, 1\}, j = 1, \ldots, 8. \end{array}$$

Los valores u que se obtengan corresponden a coeficientes a_{jt} que definen un patron de turno.

Pregunta 3 (12 puntos)
Una compañía industrial posee un conjunto de plantas en las que fabrica productos para satisfacer demandas en ciertos centros de consumo. Esta empresa quiere planificar sus operaciones sobre un horizonte consistente en T periodos. Sea F el conjunto de fábricas y sea D el conjunto de centros de demanda. Sea P el conjunto de productos. En cada centro de demanda $j \in D$ se conoce la demanda d_{jt}^k por el producto $k \in P$ en el periodo $t \in \{1, \ldots, T\}$. Cada fábrica tiene una capacidad productiva total igual a $Q_{it}, i \in F$ en el periodo t. Cada unidad del producto $k \in P$

producida en la fábrica $i \in F$ consume a_{ik} unidades de capacidad. Existen costos variables de producción, c^i_{kt} que es el costo de producir una unidad de $k \in P$ en la fábrica $i \in F$ en el periodo t y, análogamente, hay costos de inventario en cada fábrica y para cada producto, h^i_{kt}. El costo de transportar una unidad del producto k entre la fábrica i y el centro j en el periodo t es g^{tk}_{ij} y asumimos que no hay restricciones de capacidad en el transporte. El modelo que determina cuánto y cuándo producir en cada fábrica y además cuanto transportar, se muestra a continuación. Para esto se usan las variables x^i_{kt}, que es la cantidad a fabricar de $k \in P$ en $i \in F$ en el periodo $t \in 1, \ldots, T$, w^{tk}_{ij} que es la cantidad a transportar de $i \in F$ a $j \in D$ del producto $k \in P$ en el periodo $t \in 1, \ldots, T$, I^i_{kt} es el inventario:

$$
\begin{array}{lll}
\min & \sum_{t=1}^{T} \sum_{i \in F} \sum_{k \in P} \{c^i_{kt} x^i_{kt} + h^i_{kt} I^i_{kt} + \sum_{j \in D} g^{tk}_{ij} w^{tk}_{ij}\} & \\
s.t. & \sum_{k \in P} a_{ik} x^i_{kt} \leq Q_{it} & \forall i \in F, t = 1, \ldots, T \quad (a) \\
& I^i_{kt} = I^i_{k,t-1} + x^i_{kt} - \sum_{j \in D} w^{tk}_{ij} & \forall k \in P, i \in F, t = 1, \ldots, T \quad (b) \\
& \sum_{i \in F} w^{kt}_{ij} = d^k_{jt} & \forall j \in D, k \in P, t = 1, \ldots, T \quad (c) \\
& x^i_{kt} \geq 0, I^i_{kt} \geq 0, w^{tk}_{ij} \geq 0 & \forall k \in P, i \in F, j \in D, t = 1, \ldots, T \quad (d)
\end{array}
$$

donde se asume que las cantidades $I^i_{k,0}$ son conocidas (los inventarios iniciales).

a) (8 pts) Formule una relajación Lagrangeana para resolver este problema, justificando con claridad su elección de restricciones a "dualizar". Se espera que su relajación sea "buena", así que piense adecuadamente cuáles restricciones elegir para formar el dual (note que el problema es lineal). Especifique con claridad la función dual que usará e indique con precisión cómo se resuelve el correspondiente problema relajado que evalúa la función dual.

Respuesta: Hay dos posibles relajaciones naturales y ambas tienen argumentos a favor.

Consideremos primero la que corresponde a dualizar las restricciones de cota de capacidad (a). Las restricciones de cota se dualizan con multiplicadores $\lambda_{it} \geq 0$. La función dual resultante resuelve el siguiente problema relajado:

$$
\begin{array}{lll}
\min & \sum_{t=1}^{T} \sum_{i \in F} \sum_{k \in P} \{c^i_{kt} x^i_{kt} + h^i_{kt} I^i_{kt} + \sum_{j \in D} g^{tk}_{ij} w^{tk}_{ij}\} + \sum_{t=1}^{T} \sum_{i \in F} \lambda_{it} (\sum_{k \in P} a_{ik} x^i_{kt} - Q_{it}) & \\
s.t. & I^i_{kt} = I^i_{k,t-1} + x^i_{kt} - \sum_{j \in D} w^{tk}_{ij} & \forall k \in P, i \in F, t = 1, \ldots, T \quad (b) \\
& \sum_{i \in F} w^{kt}_{ij} = d^k_{jt} & \forall j \in D, k \in P, t = 1, \ldots, T \quad (c) \\
& x^i_{kt} \geq 0, I^i_{kt} \geq 0, w^{tk}_{ij} \geq 0 & \forall k \in P, i \in F, j \in D, t = 1, \ldots, T \quad (d)
\end{array}
$$

Reordenando los términos en la función objetivo tenemos:

$$
\begin{array}{lll}
\min & \sum_{t=1}^{T} \sum_{i \in F} \sum_{k \in P} \{(c^i_{kt} + \lambda_{it} a_{ik}) x^i_{kt} + h^i_{kt} I^i_{kt} + \sum_{j \in D} g^{tk}_{ij} w^{tk}_{ij}\} + \sum_{t=1}^{T} \sum_{i \in F} \lambda_{it} Q_{it} & \\
s.t. & I^i_{kt} = I^i_{k,t-1} + x^i_{kt} - \sum_{j \in D} w^{tk}_{ij} & \forall k \in P, i \in F, t = 1, \ldots, T \quad (b) \\
& \sum_{i \in F} w^{kt}_{ij} = d^k_{jt} & \forall j \in D, k \in P, t = 1, \ldots, T \quad (c) \\
& x^i_{kt} \geq 0, I^i_{kt} \geq 0, w^{tk}_{ij} \geq 0 & \forall k \in P, i \in F, j \in D, t = 1, \ldots, T \quad (d)
\end{array}
$$

Se puede ver que este problema se separa por el índice k de cada producto. Para cada producto el problema resultante es uno de producción, inventarios y flujos. De hecho, es un problema totalmente de flujo a costo mínimo el cual puede ser resulto muy fácilmente, como todos los problemas de flujo a costo mínimo. (esto puede que lo sepan muy pocos, no es obvio de inmediato y sería conocimiento "avanzado"). Esta es una buena relajación por esa razón.

La otra relajación que acá se podía hacer es dualizar las restricciones de inventario. Estas se dualizan con multiplicadores de Lagrange μ^i_{kt} sin restricción de signo. La función dual evalúa el siguiente problema:

$$
\begin{array}{llll}
\min & \sum_{t=1}^{T}\sum_{i\in F}\sum_{k\in P}\{c_{kt}^i x_{kt}^i + h_{kt}^i I_{kt}^i + \sum_{j\in D} g_{ij}^{tk} w_{ij}^{tk}\} + & \sum_{t=1}^{T}\sum_{i\in F}\sum_{k\in P}\mu_{kt}^i (I_{kt}^i - I_{k,t-1}^i - x_{kt}^i + \sum_{j\in D} w_{ij}^{tk}) & \\
s.t. & \sum_{k\in P} a_{ik}x_{kt}^i \leq Q_{it} & \forall i\in F, t=1,\ldots,T & (a)\\
 & \sum_{i\in F} w_{ij}^{kt} = d_{jt}^k & \forall j\in D, k\in P, t=1,\ldots,T & (c)\\
 & x_{kt}^i \geq 0, I_{kt}^i \geq 0, w_{ij}^{tk}\geq 0 & \forall k\in P, i\in F, j\in D, t=1,\ldots,T & (d)
\end{array}
$$

Reordenándo términos en la función objetivo tenemos:

$$
\begin{array}{llll}
\min & \sum_{t=1}^{T}\sum_{i\in F}\sum_{k\in P}\{(c_{kt}^i-\mu_{kt}^i)x_{kt}^i + (h_{kt}^i + \mu_{kt}^i - \mu_{kt-1}^i)I_{kt}^i + & \sum_{j\in D}(g_{ij}^{tk}+\mu_{kt}^i)w_{ij}^{tk}\} & \\
s.t. & \sum_{k\in P} a_{ik}x_{kt}^i \leq Q_{it} & \forall i\in F, t=1,\ldots,T & (a)\\
 & \sum_{i\in F} w_{ij}^{kt} = d_{jt}^k & \forall j\in D, k\in P, t=1,\ldots,T & (c)\\
 & x_{kt}^i \geq 0, I_{kt}^i \geq 0, w_{ij}^{tk}\geq 0 & \forall k\in P, i\in F, j\in D, t=1,\ldots,T & (d)
\end{array}
$$

Este problema se separa en varias categorías. Primero, se separa en las variables x_{kt}^i, otro distinto en las variables w_{ij}^{kt} y las variables I_{kt}^i quedan solo en la función objetivo.

Cada problema de las variables x se separa en problemas independientes por (i,t) los cuales son problemas lineales con una sóla restricción y de resolución inmediata (la variables básica es la con menor valor de $(c_{kt}^i - \mu_{kt}^i)/a_{ik}$. Algo similar pasa con los problemas en las variables w, se separan en las índices (j,t,k) y quedan problemas lineales con una sola restricción, de solución inmediata. Por otro lado, como las variables I sólo aparecen en la función objetivo, y estamos minimizando, su valor depende del coeficiente $h_{kt}^i + \mu_{kt}^i - \mu_{kt-1}$. Si este es ≥ 0, entonces $I_{kt}^i = 0$. Si es < 0, entonces el valor de la función dual es $-\infty$. Eso significa que los valores de los multiplicadores que dan esa situación no son válidos. El proceso de búsqueda del algoritmo de supgradiente deberá, entonces, tomar eso en cuenta al imponer la restricción de que $h_{kt}^i + \mu_{kt}^i - \mu_{kt-1} \geq 0$. Esta relajación puede verse como fácil de resolver también ya que separa el cálculo en problemas de fácil resolución.

b) (4 pts) Suponga que el problema dual se resuelve mediante el método de “supgradiente”. Explique cómo se obtiene, en una iteración específica, un supgradiente para la función dual construida por usted en a).

Respuesta: Una vez resuelta una evalaución de la función dual, el subgradiente corresponde a las holguras de las restricciones relajadas de cota, es decir:

$$h_{it} = Q_{it} - \sum_{k\in P} a_{ik}x_{kt}^i$$

para el caso de la primera relajación. Algo análogo para la segunda.