ICS 2121 Métodos de Optimización, Sem 2020-2
Prof. Jorge Vera

Tarea 2.

Instrucciones: Esta tarea debe realizarse en los grupos constituidos en el curso según las instrucciones dadas. La tarea debe ser respondida en forma ordenada y escrita en computador, usando LaTeXu otro procesador de texto. **No se aceptarán tareas manuscritas. De ocurrir, no serán corregidas. Debe ser subida a la plataforma de tareas a más tardar el viernes 9 de octubre de 2020 a las 23:59 horas, como un archivo PDF.** Si usan LaTeX, les daremes 2 décima de bono en la tarea. Durante su trabajo en la tarea deben respetar las normas éticas correspondientes, cualquier colaboración indebida entre los grupos sera sancionada según establecen los reglamentos de la Escuela y Universidad, pudiendo incluso implicar la reprobación del curso. Todos los temas de esta tarea pueden ser evaluados en las interrogaciones.

Los problemas 1, 2 y 3 requieren de desarrollos computacionales usando los códigos Python de Métodos de Primer Orden que se le entregaron, particularmente el de FISTA. El archivo Tarea2-2020-2.xlsx es una planilla Excel con datos de 300 personas, en donde la primera columna corresponde a mediciones de niveles de glucosa en la sangre de la persona y las 1500 columnas siguientes corresponden al nivel de expresión de diversos marcadores genéticos.

Pregunta 1 (10 pts):
Modifique los códigos Python que se entregaron para el algoritmo FISTA para que usen una matriz A y vector b suministrados externamente mediante una planilla Excel. Como en la tarea 1, le recomendamos que use paquetes que permiten interactuar con Excel, como por ejemplo `openpyxl`.
Ahora corra el set de datos suministrado y trate de obtener una solución al problema donde un alto porcentaje de coordenadas de la solución sean igual a 0, es decir, trate de identificar los marcadores que mejor explican el peso. Haga esto probando distintos valores de τ y número de iteraciones, tratando de lograr un balance adecuado entre el error en el sistema $Ax = b$ y el término de $\|x\|_1$ (el código despliega esos valores en cada iteración, con el error medido el forma relativa, vea el código). Recuerdo que, como explicamos e ilustramos en clases, las soluciones que obtendrá en esta implementación difícilmente tendrán coordenadas exactamente igual a 0, así que tendrá que decidir qué es lo que "redondea" a 0. Muestre gráficos con la convergencia y explique con claridad cómo llegó a su solución final.

Pregunta 2 (10 pts):
Ahora queremos usar el algoritmo de Frank-Wolfe para el mismo problema. Como usted sabe, el problema que queremos resolver es el de regularización L1:

$$\min_{x\in\mathbb{R}^n} \tau\|x\|_1 + \|Ax - b\|_2^2$$

En clases mostramos que este problema podía ser aproximado, a su vez, por el problema:

$$\begin{array}{ll} \min & \|Ax - b\|_2^2 \\ s.a. & \|x\|_1 \leq \rho \end{array}$$

para algún valor de ρ adecuado.
Escriba el detalle del algoritmo de Frank-Wolfe para el caso específico de este problema. Puede usar como base el pseudocódigo que se mostró en clases, pero haga explícito el problema de Programación Lineal que se debe resolver en cada iteración.

Pregunta 3 (10 pts):

En este problema tendrá que implementar y ejecutar el algoritmo de Frank-Wolfe para los mismos datos del Problema 1. Le sugerimos que use como punto de partida el mismo código que usó para FISTA, pero modifíquelo para poder implementar el algoritmo de Frank-Wolfe. Necesitará hacer un cambio bastante grande al código pero hay algunas partes que le servirán, como el loop general y parte del cálculo del gradiente. También tendrá que implementar la resolución del problema de Programación Lineal que es parte del método de Frank-Wolfe. Para esto último, use las rutinas de Gurobi en Python. Si no tiene aún instalado Gurobi, hágalo según las instrucciones que se le enviaron a través de los avisos del curso.
Ahora, ejecute el algoritmo en el mismo problema de la Pregunta 1, probando distintos valores de ρ. Muestre gráficos con la convergencia y explique con claridad cómo llegó a su solución final. ¿Cómo se compara el funcionamiento y las soluciones que pudo obtener con las que alcanzó mediante FISTA? Discuta sobre el comportamiento de los métodos.

Pregunta 4 (10 pts):
Esta pregunta requiere que usted explore referencias relevantes a los temas estudiados en el curso. En clases planteamos que los Métodos de Primer Orden tienen gran importancia en aplicaciones de "Ciencias de Datos", "Big Data", etc., particularmente la técnica LASSO, hace uso de esto. Busque en la web un artículo publicado en alguna revista científica y que muestre uso de la técnica LASSO (o Regularización L1 u otro de los nombres que hemos usado en el curso). Elija un área de aplicación, para buscar, de la siguiente lista:

1. Estadísticas y Econometría aplicada en problemas de gestión y económicos.
2. Investigación en temas médicos, biología o similares.
3. Reconstrucción de Imágenes.
4. Física experimental, astronomía, astrofísica y relacionados.

Del artículo elegido por ustedes, escriban un resumen de máximo dos páginas. Este resumen no puede ser simplemente copia del "abstract" del artículo sino que debe incluir su propio análisis, destacar los aportes del artículo, la relevancia de la aplicación, si acaso se discute la metodología de optimización subyacente, etc. Debe incluir también una opinión personal (es decir, del grupo, se entiende) sobre el artículo, refiéranse, entre otras cosas, a aplicaciones o caminos futuros que puede tener el desarrollo del paper que eligieron.

Pregunta 5 (10 pts):
Usted debe recordar el problema clásico de planificación de producción, el que reproducimos aquí: consideramos n productos y m recursos, en un horizonte de T meses. Los recursos tienen disponibilidades máximas y se debe satisfacer demanda. Los siguientes son los parámetros del problema:

d_{jt}**:** demanda por producto j en periodo t, $j = 1, ..., n, t = 1, ..., T$.

b_{it}**:** Disponibilidad del recurso i en periodo t, $i = 1, ..., m, t = 1, ..., T$.

c_{jt}**:** Costo unitario de producción de j en t, $j = 1, ..., n, t = 1, ..., T$.

a_{ij}**:** Cantidad de recurso i usado en producir una unidad de producto j, $i = 1, ..., m, j = 1, ..., n$.

h_{jt}**:** Costo de dejar una unidad de producto j en inventario en el período t, $j = 1, ..., n$, $t = 1, ..., T$

El modelo del problema usa dos tipos de variables: x_{jt}, que es la producción de j en t, y se usa otra variable, I_{jt}, para registrar el inventario. El modelo es:

$$
\begin{array}{lll}
\min & \sum_{t=1}^{T} \sum_{j=1}^{n} \{c_{jt} x_{jt} + h_{jt} I_{jt}\} & \\
s.t. & I_{jt} = I_{jt-1} + x_{jt} - d_{jt}, & t = 1, \ldots, T; j = 1, \ldots, n \\
& \sum_{j=1}^{n} a_{ij} x_{jt} \leq b_{it}, & i = 1, \ldots, m; t = 1, \ldots, T \\
& x_{jt} \geq 0, I_{jt} \geq 0, & t = 1, \ldots, T; j = 1, \ldots, n
\end{array}
$$

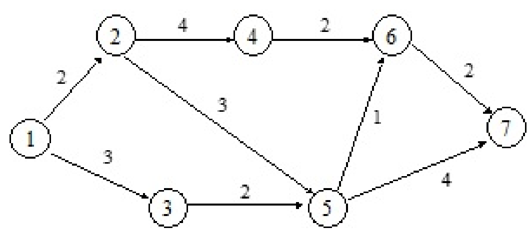
(los I_{j0} son datos)
Escriba el dual de este problema y entregue una interpretación de las distintas variables duales.

Pregunta 6 (10 pts):
Considere un problema de Programación Lineal en forma estándar $\min\{c^T x : Ax = b, x \geq 0\}$, donde $A \in \mathbb{R}^{m\times n}$, $b \in \mathbb{R}^m$, $c \in \mathbb{R}^n$. Suponga que $b \geq 0$. Como usted sabe, el Simplex comienza con el problema de Fase I, que tiene la forma:

$$\begin{aligned} w^* = \min \quad & e^T v \\ s.a. \quad & Ax + Iv = b \\ & (x, v) \geq 0 \end{aligned}$$

donde $v \in \mathbb{R}^m$ es el vector de variables artificiales y $e \in \mathbb{R}^m$ es el vector cuyas componentes son todas iguales a 1. Suponga que el valor óptimo de este problema de Fase I es positivo, es decir, $w^* > 0$. Construya el dual de este problema de Fase I y deduzca, a partir de ahí, que el problema original tiene que ser infactible. Justifique con claridad sus argumentos.

Pregunta 7 (10 pts):
Un proyecto se ha dividido en actividades que deben ser realizadas respetando relaciones de precedencia entre ellas. Se ha decidido utilizar una red dirigida para representar estas relaciones, en la cual cada arco está asociado a una actividad, y cada nodo representa el término o inicio de las actividades que llegan o salen de él, respectivamente. Por ejemplo, en la red de figura, las actividades (1,2) y (1,3) son las que se deben realizar al inicio del proyecto. El nodo 1 representa el evento o "hito" de inicio. Por otro lado, el nodo 5 representa el término de las actividades (3,5) y (2,5), y el comienzo de las actividades (5,6) y (5,7). Estas dos últimas no se pueden realizar mientras no se hayan completado las actividades (2,5) y (3,5). Se ha estimado la duración de cada actividad, denotado por d_{ij} y este dato se muestra igualmente en la figura:



Más formalmente y en general, un proyecto se puede, entonces, representar con un grafo dirigido $G = (N, A)$ donde un nodo r representa el inicio y un nodo s representa el final del proyecto. Los arcos son las actividades y los nodos son inicios y términos de conjuntos de actividades, d_{ij} es la duración de la actividad (i, j).

a) (5 pts) Denotemos por t_i la variable de decisión que representa el instante de tiempo en que pueden comenzar las actividades cuyos arcos salen del nodo i. Plantee un modelo general de programación lineal que minimice la duración total del proyecto representado por un grafo G, respetando las relaciones de precedencia. (Ind: piense en cómo representar con estas variables lo que se debe cumplir en los tiempos con las relaciones de precedencia, puede apoyarse en el grafo ejemplo, pero tiene que construir un modelo general, que sirva para cualquier proyecto.)

b) (5 pts) Formule el problema dual asociado al modelo lineal anterior y utilizando las relaciones de dualidad, muestre que la ruta *más larga* entre el nodo r y el nodo s, esto es la secuencia de actividades que tienen la mayor duración total, corresponde a la duración mínima del proyecto. (Esta ruta se conoce como *ruta crítica* en la jerga de planificación de proyectos).

Pregunta 8 (10 pts):
Suponga que dispone de n instrumentos de inversión (bonos, por ejemplo). Sean $p_1, \ldots, p_n$ los precios de estos bonos al comienzo del periodo de inversión. Al final del horizonte de inversión, el valor de los bonos es $v_1, \ldots, v_n$. Supongamos que se invierte $x_1, \ldots, x_n$ en cada uno de los bonos (se puede, eventualmente, tener $x_j < 0$, es lo que en finanzas se llama una "posición corta"). Entonces el valor de la inversión inicial es

$\sum_{j=1}^{n} p_j x_j$ y el retorno final es $\sum_{j=1}^{n} v_j x_j$. Dado que el retorno real es incierto, es habitual asumir algunos escenarios futuros. Específicamente, supongamos que se asume que pueden ucurrir cualesquiera de m escenarios de retornos, $v^1, \ldots, v^m$, donde $v^i \in \mathbb{R}^n$ y corresponde a un vector de n componentes, donde v_j^i es el retorno del instrumento j en el escenario i. Si ocurre el escenario i, el retorno será $\sum_{j=1}^{n} v_j^i x_j$. Si existe un vector de inversiones x tal que $\sum_{j=1}^{n} p_j x_j < 0$ y $\sum_{j=1}^{n} v_j^i x_j \geq 0$, para todos escenario $i = 1, \ldots, m$, se dice que hay una "oportunidad de arbitraje", que significa que hay una garantía de obtener utilidades en la inversión, libre de riesgos. La teoría económica-financiera impone habitualmente la condición de que no exista arbitraje. Muestre que para que no exista arbitraje, entonces deben existir valores $u_1 \geq 0, u_2 \geq 0, \ldots, u_m \geq 0$ tales que $p_j = \sum_{i=1}^{m} v_j^i u_i$. Trate de interpretar el resultado, desde un punto de vista económico.