Pontificia Universidad Católica de Chile
Escuela de Ingeniería
Departamento de Ingeniería Industrial y de Sistemas
Profesores: Gustavo Angulo, Raimundo Cuadrado,
Camila Balbontin, Gonzalo Pérez
Ayudante Coordinador de Tareas: Juan Enrique Hurtado

ICS1113 - Optimización
1er semestre del 2024

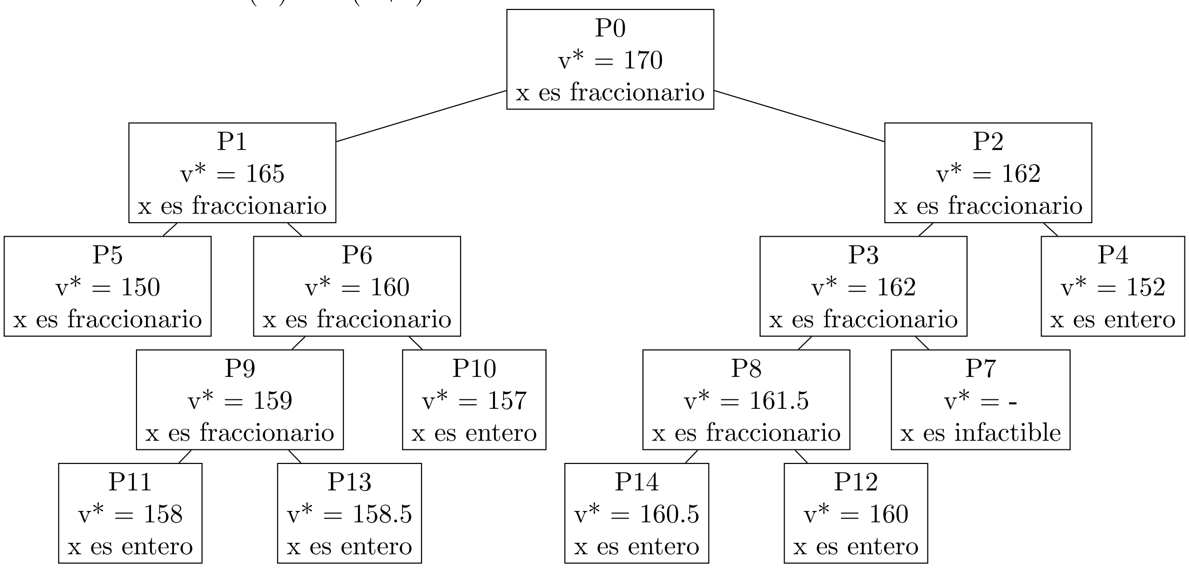
Tarea 3 ICS1113

Reglas de la tarea:

- La tarea se puede realizar de a pares o de forma individual (las parejas pueden ser de diferentes secciones). En caso de ser de a pares se debe inscribir con su pareja en los grupos de Canvas.
- La tarea debe ser subida **por solo uno de los integrantes** a Canvas.
- La tarea se entrega el día 26/06/2024 a las 23:59 horas.
- Existirá la opción de usar un cupón de atraso solamente una vez entre las tres tareas, dando la posibilidad de entregar hasta 24 horas después de finalizada la entrega. Para poder acceder a este cupón se deben inscribir con el nombre de los integrantes de la tarea en un Google Forms que será habilitado en su debido tiempo y que cierra junto a la fecha oficial de la entrega. En caso de ser de a pares ambos deben tener un cupón disponible y se usará uno por cada uno y tendrán 24 horas extra (no 48).
- Las tareas atrasadas no serán corregidas, así que recuerden entregarlas a tiempo.
- Esta tarea es grupal y el desarrollo y discusión debe ocurrir dentro de cada grupo. No se distribuyan la resolución de las preguntas por separado, hagan realmente un trabajo grupal de desarrollo ya que el no hacerlo va contra la idea de aprendizaje colaborativo y preparación individual de la interrogación a la que está asociada la tarea. Pueden discutir los problemas con los profesores y los ayudantes del curso, pero al final cada grupo debe entregar sus propias soluciones, desarrolladas y escritas por el grupo. La copia o intento de copia a otros grupos será sancionada dependiendo la gravedad (a ser determinada por el equipo docente del curso) con consecuencias que podrían ir desde un 1.0 en la nota de la Tarea hasta la escalación a la Dirección de Docencia de la Escuela, con posible **REPROBACIÓN AUTOMÁTICA** del curso.
- Por entregar en latex se darán 3 décimas para la tarea entregada, las décimas no son transferibles a otras actividades. Para subirlo tendrá que enviar un archivo .pdf y el archivo .tex correspondientes (también se puede incluir el .zip de Overleaf). En caso de no entregar en latex basta con subir un archivo .pdf que contenga las respuestas. Si se hace a mano, debe ser una tarea legible o habrán descuentos a criterio del corrector. En ambos casos los .pdf tienen que contener los nombres de los integrantes.
- Se abrirá un foro en Canvas donde se responderán las dudas que puedan surgir del enunciado.

Pregunta 1 **[20 Puntos]**

1. Considere el siguiente árbol completo del algoritmo Branch&Bound para la resolución de un problema de optimización entero. Un algoritmo computacional fue resolviendo cada uno de los subproblemas en el orden P(N) $\rightarrow$ P(N+1).



   - Indique si es que es un problema de maximización o de minimización. Explique.
   - Elabore una tabla en cuyas columnas estén: nombre de subproblema resuelto, valor incumbente después de resolver ese subproblema, mejor cota después de resolver ese subproblema y el gap. En cada fila de la tabla, deben ir ordenados de manera creciente los 15 subproblemas y su información. En la tabla debe incluir el $+\infty$ o el $-\infty$ cuando corresponda. Considere el gap como gap = $\|$incumbente - mejor cota$\|$/incumbente.
   - Mencione el/los nodo(s) que no se debió/debieron resolver si se hubiera ejecutado el algoritmo B&B de manera correcta.
   - Explicite el nombre del primer nodo en que se podría detener el algoritmo con un gap menor al 2,5 %.

2. Dados $a \in \mathbb{R}^n_+$ y $b \in \mathbb{R}_+$, sea $S = \{x \in \{0,1\}^n : \ a^\top x \geq b\}$. Decimos que $E \subseteq N = \{1, \ldots, n\}$ es un empaquetamiento si $\sum_{j \in E} a_j < b$. Demuestre que la desigualdad de empaquetamiento $\sum_{j \in N \setminus E} x_j \geq 1$ es válida para S.

3. Sea $D \subseteq \mathbb{R}^2$ definido como $D = \{(x_1, x_2) \in \mathbb{Z} : \ -x_1 + 5x_2 \geq 16; \quad x_1 + x_2 \geq 12; \quad 2{,}6x_1 + x_2 \leq 43{,}2; \quad x_2 \leq 12; \quad -9x_1 + 5x_2 \leq -3\}$. Sea $c = (1, a) \in \mathbb{R}^2$ un vector. Determine para qué intervalo(s) abierto(s) del parámetro a el valor óptimo de $\min\{c^T x : \quad x \in D\}$ es igual al valor óptimo de su relajación lineal. Justifique.

4. Se tiene el siguiente problema de programación lineal entera.

$$
\begin{aligned}
\text{P) max} \quad & 4x_1 + 3x_2 + 2x_3 + 4x_4 + x_5 + x_6 \\
\text{s.a.} \quad & 6x_1 + 2x_2 + 7x_3 + 2x_6 \leq 10 \\
& 7x_1 + 3x_2 + 5x_4 + 15x_5 \leq 14 \\
& x_1, x_2, x_3, x_4, x_5, x_6 \in \{0,1\}
\end{aligned}
$$

   Encuentre 4 subconjuntos covers distintos que se pueden deducir de las restricciones, y escriba el nuevo problema con las 4 nuevas restricciones añadidas.

Pregunta 2 **[20 Puntos]**

1. Indique la veracidad de cada una de las siguientes afirmaciones. Justifique en caso de ser verdad (V) o de un contraejemplo en caso de ser falsa (F). Solo se asignará puntaje con una correcta justificación o contraejemplo.

   a) Sea P un problema de optimización no lineal cuya solución óptima activa dos restricciones. Es posible que el problema Q, idéntico a P, pero sin sus restricciones tenga su misma solución óptima.

   b) Es posible que exista una función que sea cóncava y convexa a la vez en una parte de su dominio.

   c) En un problema de minimización no lineal irrestricto que admite solución óptima siempre hay una mayor cantidad de mínimos locales que de máximos locales.

   d) En un problema de optimización no lineal, en todo punto que cumpla las condiciones de KKT se tiene que las variables de las restricciones activas λ y μ son mayores que cero.

   e) Todo problema de minimización con restricciones, que tenga una función objetivo con matriz hessiana definida positiva en todo el dominio del problema, tiene solución única.

   f) En todo punto singular se tiene al menos un par de gradientes de las restricciones activas siendo linealmente dependientes.

   g) Todo punto del dominio de un problema no lineal cumple las condiciones de KKT o es un punto singular.

   h) En un punto singular el gradiente de la f.o. es siempre linealmente independeinte del conjunto de gradientes de las restricciones activas en dicho punto.

2. Considere el siguiente problema de optimización.

$$(P) \quad \begin{array}{l} \min f(x) \\ \text{s.a } g_i(x) \leq 0 \quad \forall i = 1, \ldots, m \end{array}$$

   Sea $\bar{x} \in \mathbb{R}^n$ un mínimo local de (P). Sea $I = \{i : g_i(\bar{x}) = 0\}$ el conjunto de los índices de las restricciones activas en $\bar{x}$. Suponga que $f, g_i \in C^1(\mathbb{R}^n)$ $\forall i = 1, \ldots, m$. Demuestre que toda dirección de descenso $d : \nabla f(\bar{x})^T d < 0$ va en una dirección en que se viola al menos una de las restricciones de I, es decir, existe $i \in I$ tal que $\nabla g_i(\bar{x})^T d > 0$. Use las ecuaciones de KKT.

3. Se tiene el siguiente problema de optimización no lineal:

$$\begin{array}{ll} \text{máx} & (x_2 - 3)^2 - x_1 \\ \text{s.a.} & x_1^2 + x_2^2 \leq 9 \\ & x_1^2 + (x_2 - 4)^2 \leq 4 \end{array}$$

   a) Grafique el problema y encuentre la solución y el valor óptimo.

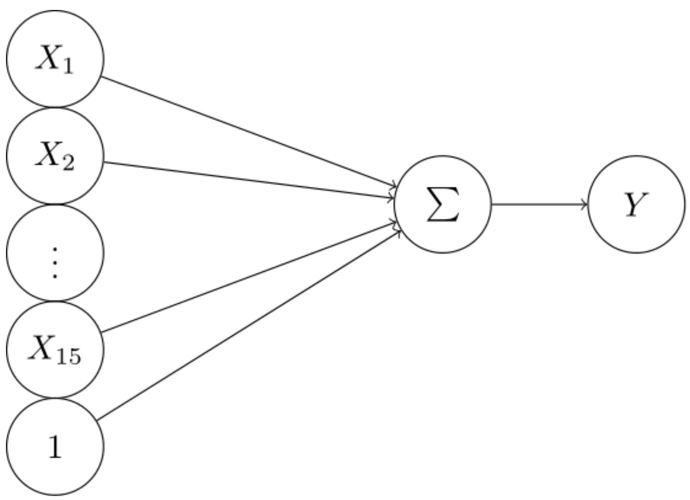
   b) Verifique las condiciones de KKT para el punto encontrado en (a). Resuelva el sistema de ecuaciones.

Pregunta 3 **[20 Puntos]**

1. Optimización en Machine Learning

1.1. Descripción del problema

La empresa Cabo de Hornos, es de las principales productoras de almohadas de alta gama. La empresa le vende a cada cliente una almohada con cierta cantidad de relleno (Y) en gramos de plumas. Cada persona, por sus características físicas, tiene sólo una almohada que es máximamente cómoda por su cantidad de relleno. La empresa gasta muchos recursos en tiempo cuando los clientes se devuelven a cambiar el relleno de la almohada, que se debe a que la cantidad de relleno que se le recomendó al cliente inicialmente no era totalmente cómodo. Es por lo anterior que Cabo de Hornos ha decidido contratarlo a usted para que con sus conocimientos de optimización elabore y resuelva un modelo matemático que permita entregarle a sus clientes una recomendación óptima de relleno de almohada basado en ciertas mediciones de sus dimensiones corporales.
Basado en lo anterior, la empresa ha decidido hacer una encuesta pagada a 100 de sus clientes que están completamente satisfechos con su almohada. Se les pide que tomen quince medidas de sus características corporales, las cuales son: peso (x_1), estatura (x_2), edad (x_3), contorno de la cabeza (x_4), distancia hombro-cuello (x_5), altura del cuello(x_5),.. etc. También se les pide que indiquen los gramos de plumas de su almohada (y).
El objetivo de este ejercicio de machine learning es construir una función Y = f(X) que permita entregar una recomendación Y para cada persona basado en sus medidas corporales. Para esto se entrenará una red neuronal del tipo regresión lineal usando la información de las encuestas. El modelo es el siguiente:



Este diagrama del modelo a utilizar explica cómo es el recorrido de transformaciones para pasar desde los inputs X al output Y. Todos y cada uno de los arcos (menos el final) tienen asociados unos valores denominados pesos. Llámese w_i a los pesos de los arcos que conectan los inputs x_i con el nodo de la sumatoria. Llámese además b al peso que conecta la constante 1 a la sumatoria final. Para elaborar la función deseada se debe conseguir el objetivo de asignarle valores a todos los pesos descritos anteriormente. Pero no pueden ser cualquier valor, pues debieran representar de buena manera la función que conecta X con Y. Si, por ejemplo, todos los pesos tuvieran valor cero, el modelo siempre predeciría erróneamente que la almohada óptima es de Y = 0 gramos, para cualquier persona de características X. Esto es un problema porque va en contradicción con los datos recopilados, ya que ningún relleno de almohada en la encuesta es de valor cero, sino que son mayores. La idea es encontrar pesos que minimicen la diferencia entre lo predicho por el modelo para una entrada X y el verdadero output Y de esa entrada X.

1.2. Formulación Matemática 1

El primer problema busca encontrar los pesos w_i y la constante b del modelo a entrenar descrito anteriormente. Los parámetros de este problema serán los valores X e Y, los cuales se obtienen de la encuesta. Sea $r \in \{1, ..., 100\}$ el subíndice para cada una de las muestras de clientes que respondieron la encuesta.

Se deberá resolver el siguiente problema de optimización no lineal irrestricto para encontrar los pesos y constante óptimos:

$$\min_{w,b} \quad \sum_{r=1}^{100} (b + \sum_{i=1}^{15} w_i \cdot x_{i,r} - y_r)^2 \tag{1}$$

Parte 1

Se le pide que elabore un código main1.py y explicite en su consola los valores de cada uno de los pesos w_i y b para cada uno de las características. Debe mencionar el nombre de la característica y su respectivo ponderador en orden. Recuerde también imprimir en consola el valor del término b (es una constante en gramos).

1.3. Formulación Matemática 2

Después de resolver el modelo anterior, la empresa se dio cuenta de que era mucho tomarle 15 medidas a cada cliente para determinar la almohada óptima a usar. Pero... ¿Qué características usar? ¿Cuáles son más importantes? Está claro que los pesos w_i tienen relación con la respuesta. Se propone agregar un factor λ que penalice la existencia de valores w_i (ya sea positivos o negativos). Se definen variables z_i que representen la norma 1 (valor absoluto) de las variables w_i.

$$\min_{w,z,b} \quad \sum_{r=1}^{100} (b + \sum_{i=1}^{15} w_i \cdot x_{i,r} - y_r)^2 + \lambda \cdot \sum_{i=1}^{15} z_i \tag{2}$$

$$-z_i \leq w_i \quad \forall i \in \{1, ..., 15\} \tag{3}$$

$$w_i \leq z_i \quad \forall i \in \{1, ..., 15\} \tag{4}$$

Parte 2

Se le pide que elabore un código main2.py en que implemente la formulación 2. El parámetro λ es un regularizador en la regresión lineal. Éste permite castigar la magnitud de los valores de w_i si es que no está realizando cambios significativos en minimizar el error cuadrático, y si son fácilmete reemplazables por los otros pesos. Note que si $\lambda = 0$, no hay castigo y si $\lambda = \infty$ todos los w_i=0. En su código debe ir aumentando progresivamente los valores de λ, y viendo los valores de los pesos w_i, hasta que sean distintos de cero exactamente 5 de éstos últimos. Esos representarán las características más significativas que determinan el relleno óptimo en una regresión lineal, según los datos. Determine el conjunto $E \subseteq \{1, ..., 15\}$ de 5 características más importantes y que su código lo imprima en pantalla al correr. Imprima en pantalla el valor λ utilizado para encontrar E. Considere que, en Python, el cero a veces se expresa como una potencia de 10 de exponente muy negativo (-11 o incluso menor).

Parte 3

Se le pide que elabore un código main3.py en que implemente de nuevo la formulación 1 del problema (sin el λ ni z). Sin embargo, esta vez, debe modificar la formulación para incluir solamente aquellas características que resultaron ser mejores predictoras en el modelo, y que estén en el conjunto E de la parte 2. Note que el resultado de la parte 3 es la que realmente le importa a la empresa Cabo de Hornos para predecir el relleno óptimo. Después de resolver el modelo, su código debe imprimir en la consola un string que represente el modelo final de la regresión que usted le entregará a la empresa.
Debe imprimirse siguiendo un formato:

$$Y = \sum_{i \in E} w_i \cdot x_i + b \tag{5}$$

Y mencionando qué representa cada uno de los 5 x_i, qué representa la variable Y, y la unidad de medida en que están cada uno.

1.4. Base de datos:

Junto a la tarea se encontrará un archivo.csv que tiene los datos recopilados en la encuesta. El formato de los datos es tal que cada fila representa la información de un usuario y en la primera fila está la información de lo que representa cada columna. Ese es el archivo con que se verificará su código.