Privacy Preserving Linear Programming

Mansi Goel

Department of Mathematics
School of Natural Sciences
Shiv Nadar University

May 10, 2021

SHIV NADAR UNIVERSITY

Table of contents

SHIV NADAR UNIVERSITY

What is Privacy Preserving Linear Programming?

With the rapid increase in computing, storage, and networking resources, data is not only collected and stored, but also analyzed. This creates a serious privacy problem which often inhibits the use of this data.

In this presentation, we investigate the solving of a fundamental optimization problem — a linear program which is formulated by combining the data of p entities which are unwilling to make their data public.

What is Privacy Preserving Linear Programming?

The linear program using a random linear transformation that will:

- Not reveal any of the privately held data.
- Give a publicly available exact minimum value to the original linear program.
- Components of the solution vector corresponding to each entity can be decoded only by the entity and revealed only if the owners agree.

Examples

$\geq ?$

Yao's Millionaires' Problem

How can Alice & Bob determine who is richer
without revealing their wealth?

$a \geq ? b$

Analogous Problem

How can we determine $a \geq b$
without revealing a & b?

Examples

Royal Dutch Shell and Cargill wish to transport Crude Oil and Cotton respectively with the common goal of minimizing cost of transportation. They do not wish to reveal their data such as amount of crude oil/cotton which can be transported in one tanker, the restrictions on transport quantity etc.

Notations

- All vectors are column vectors unless transposed to a row vector by a prime '.
- $A \in \mathbb{R}^{m \times n}$ signifies a real $m \times n$ matrix.
- For a vector $x \in \mathbb{R}^n$, x_j signifies the j-th component or j-th block of components.
- For a matrix A, A_i signifies the i-th component or block of components.

Formulating the Problem

We consider the linear program,

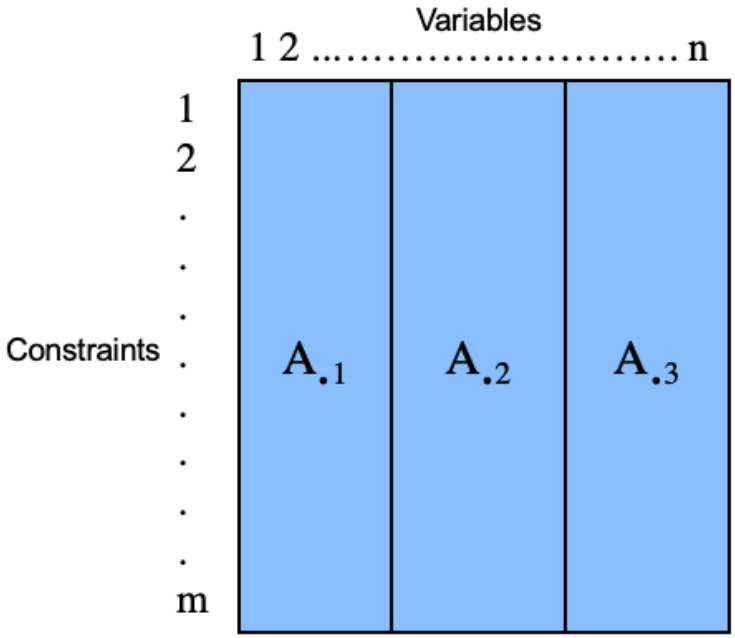
$$\min_{x \in X} c'x \text{ where } X = \{x \mid Ax \geq b\}, \tag{1}$$

Where, $A \in \mathbb{R}^{m \times n}$, the cost vector $c \in \mathbb{R}^n$ and X represents the feasible region.

The matrix A along with the cost vector c, that is $\begin{bmatrix} c' \\ A \end{bmatrix}$ is divided into **p** **vertical blocks**, each blocked own by a distinct entity that is unwilling to make it's block of data public.

Each block has $(m+1)$ rows and n_i, $i = 1, ..., p$ columns with $n_1 + n_2 + ... + n_p = n$.

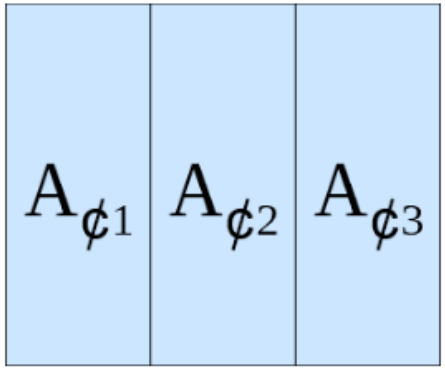
Vertically Partitioned Data



Each entity holds all the constraints but only some variables out of all the n variables.

For instance, in the example shown above, Royal Dutch Shell and Cargill hold all the transportation constraints but Royal Dutch Shell holds the variable Crude Oil and Cargill holds the variable Cotton.

Vertically Partitioned Data



SHIV NADAR UNIVERSITY

Problem Transformation

We wish to solve (1) without revealing any privately held data. We shall achieve this by proceeding as follows :

Each of the p entities chooses its own privately held random matrix $B_{.j} \in \mathbb{R}^{k \times n_j}$, $j = 1, ..., p$ where $k \geq n$. A value of k is agreed upon by all the p entities such that $k \geq n$. Define :

$$B = [B_{.1} B_{.2} \ldots\ldots B_{.p}] \in \mathbb{R}^{k \times n} \tag{2}$$

Since B is a random matrix, its rank is n [2]. We choose $k \geq n$ because we require rank of B as n to be able to define a transformation which will help formulate the secure LP.

Note : A random matrix is a matrix wherein each entry is a random variable. It has full rank with probability 1 i.e for a random matrix $A \in \mathbb{R}^{m \times n}$, rank(A) = min{m,n}.

Problem Transformation

We now define the following transformation which takes x to u :

$$x = B'u \tag{3}$$

Here, B' is a rectangular matrix with rank as column rank so using Moore - Penrose inverse,
(3) $\implies$

$$u = B\left(B'B\right)^{-1}x \tag{4}$$

We now transform our original linear program into the following **secure linear program** :

$$\min_{u \in U} c'B'u \text{ where } U = \left\{u \mid AB'u \geq b\right\} \tag{5}$$

Moore Penrose Inverse

Moore Penrose inverse A^+ of a matrix A is a generalized inverse/pseudoinverse.

Definition

For $A \in \mathbb{R}^{m \times n}$, a pseudoinverse of A is defined as a matrix $A^+ \in \mathbb{R}^{m \times n}$ satisfying:

- $AA^+A = A$
- $A^+AA^+ = A^+$
- $(AA^+)^T = AA^+$
- $(A^+A)^T = A^+A$

Moore Penrose Inverse

A^+ exists for any matrix A but when A has full rank i.e rank = min{m,n}, then A^+ can be written as a simple formula:

- When A has full column rank, $A^+ = (A^T A)^{-1} A^T$. This constitutes a left inverse since, $A^+ A = I$.
- When A has full row rank, $A^+ = A^T (AA^T)^{-1}$. This constitutes a right inverse since, $AA^+ = I$.

Problem Transformation

We refer to (5) as a secure LP since it does not reveal any of the privately held data $\begin{bmatrix} c_j' \\ A_{.j} \end{bmatrix}$.

This is so because it's impossible to compute c_j and A_j from the revealed products $c_j' B_{.j}'$ and $A_{.j} B_{.j}'$ respectively without knowing $B_{.j}$.

We now relate our original linear program (1) to the transformed linear program (5) in a way such that we are able to obtain a solution to (1) by solving (5). The following proposition helps us achieve that.

Proposition

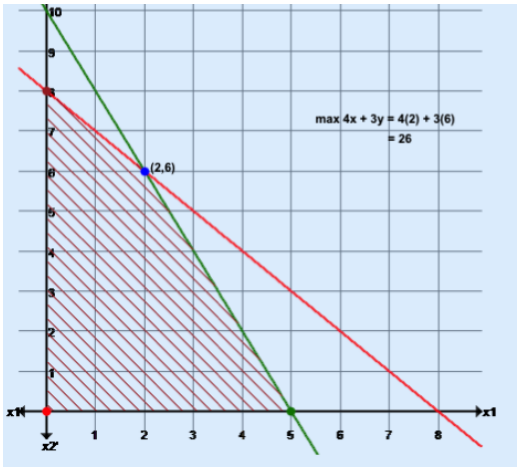
Let $k \geq n$ for the random matrix $B \in \mathbb{R}^{k \times n}$ of (2). The secure linear program (5) is solvable (optimal solution exists) if and only if the linear program (1) is solvable in which case the extrema of both linear programs are equal.

Primal and Dual Linear Programs

For any linear program (LP), there is a closely related LP called the **dual**. The feasible and optimal solutions of the dual provide useful information about the original LP which we refer to as the **primal**.

Consider the LP ,

$$\begin{aligned} \max \quad & 4x + 3y \\ \text{s.t.} \quad & x + y \leq 8 \\ & 2x + y \leq 10 \\ & x, y \geq 0 \end{aligned}$$

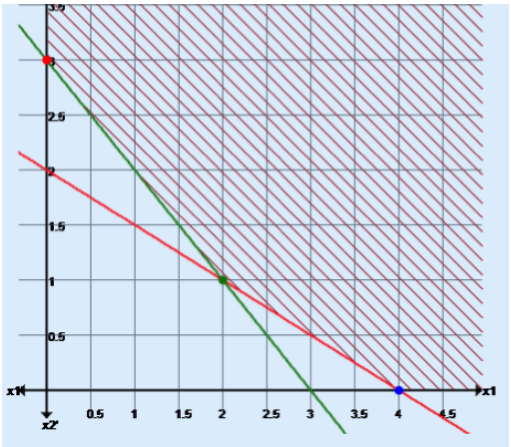


Graphically, the optimal solution of the above LP is (2,6) and the objective function value is 26.

Primal and Dual Linear Programs

Given the above LP, we introduce another LP like,

$$\begin{aligned} \min \quad & 8a + 10b \\ \text{s.t.} \quad & a + 2b \geq 4 \\ & a + b \geq 3 \\ & a, b \geq 0 \end{aligned}$$



Graphically, the optimal solution of the above LP is (2,1) and the objective function value is 26.

Construction of Dual

Consider the given LP called the **primal** where $x \in \mathbb{R}^n$, $c \in \mathbb{R}^n$, $b \in \mathbb{R}^m$, $A \in \mathbb{R}^{m \times n}$.

$$\begin{aligned} \max_{x} \quad & c^T x \\ \text{s.t.} \quad & Ax \leq b \\ & x \geq 0 \end{aligned} \tag{1}$$

Given (1), we construct another LP called the **dual**;

$$\begin{aligned} \min_{w} \quad & b^T w \\ \text{s.t.} \quad & A^T w \geq c \\ & w \geq 0 \end{aligned} \tag{2}$$

These LPs together are called the **primal - dual pair**.

- No.of dual constraints = No.of primal variables
- No.of dual variables = No.of primal constraints

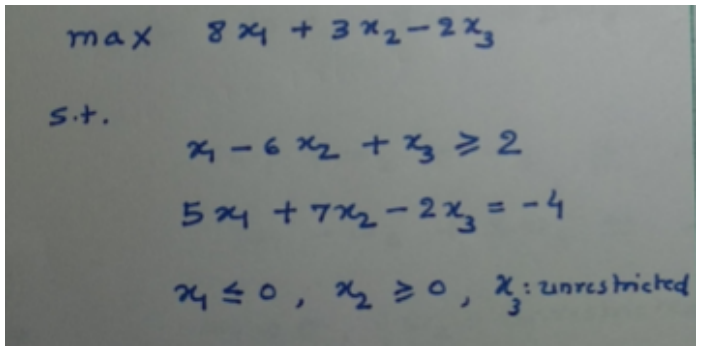
SHIV NADAR UNIVERSITY

Constructing of Dual

To decide the equalities and inequalities in the dual we use the following rule table;

	MINIMIZATION PROBLEM		MAXIMIZATION PROBLEM	
variables	≥ 0	$\Leftrightarrow$	$\leq$	constraints
	≤ 0	$\Leftrightarrow$	$\geq$	
	unrestricted	$\Leftrightarrow$	$=$	
constraints	$\geq$	$\Leftrightarrow$	≥ 0	variables
	$\leq$	$\Leftrightarrow$	≤ 0	
	$=$	$\Leftrightarrow$	unrestricted	

For example,



SHIV NADAR UNIVERSITY

Important Theorems

Theorem 1 - Weak Duality

Let x be feasible for (1) and w be feasible for (2), then $c^T x \leq b^T w$.

Proof: Since x and w are feasible for (1) and (2) we have
$Ax \leq b$, $x \geq 0$
and $A^T w \geq c$, $w \geq 0$
$Ax \leq b \Rightarrow x^T A^T \leq b^T \Rightarrow x^T A^T w \leq b^T w$
$A^T w \geq c \Rightarrow w^T A \geq c^T \Rightarrow w^T Ax \geq c^T x$
Since $x^T A^T w$ is a scalar, $x^T A^T w = w^T Ax$
Therefore, $c^T x \leq w^T Ax \leq b^T w$

Corollary 1

Let $\bar{x}$ be feasible for (1) and $\bar{w}$ be feasible for (2). Also let $c^T \bar{x} = b^T \bar{w}$.
Then, $\bar{x}$ is optimal to (1) and $\bar{w}$ is optimal to (2).

SHIV NADAR UNIVERSITY

Important Theorems

Theorem 2 - Strong Duality

- Let $\bar{x}$ be an optimal solution of (1). Then $\exists$ a $\bar{w}$ which is optimal to (2) and $c^T\bar{x} = b^T\bar{w}$.
- Let w^* be an optimal solution of (2). Then $\exists$ a x^* which is optimal to (1) and $c^T x^* = b^T w^*$.

Theorem 3 - Existence Theorem

Exactly one of the following statements are true:

- If (1) and (2) both have feasible solutions then both have optimal solutions.
- If (1) (or (2)) has an unbounded solutions, then (2) (or (1)) is infeasible.
- If (1) (or (2)) is infeasible but (2) (or (1)) is feasible then (2) will have unbounded solution.

SHIV NADAR UNIVERSITY

Proposition

Proposition

Let $k \geq n$ for the random matrix $B \in \mathbb{R}^{k \times n}$ of (2). The secure linear program (5) is solvable (optimal solution exists) if and only if the linear program (1) is solvable in which case the extrema of both linear programs are equal.

The original LP :

$$\min_{x \in X} c'x \text{ where } X = \{x \mid Ax \geq b\}, \tag{1}$$

Dual of (1) :

$$\max_{v \in V} b'v \text{ where } V = \left\{v \mid A'v = c, v \geq 0\right\} \tag{6}$$

SHIV NADAR UNIVERSITY

Proposition

The transformed LP :

$$\min_{u \in U} c' B' u \text{ where } U = \{u \mid AB'u \geq b\} \tag{5}$$

Dual of (5) :

$$\max_{w \in W} b' w \text{ where } W = \{w \mid BA'w = Bc, w \geq 0\} \tag{7}$$

So, we now have two primal-dual pairs namely (1)-(6) and (5)-(7).

The structure of the proof is as follows :

1. (1) - (6) solvable => (5) - (7) solvable.
2. (5) - (7) solvable => (1) - (6) solvable.
3. Optimal solution of (1) and (5) are the same.

Proof (Part 1)

($\Rightarrow$)
Let $\bar{x}$ and $\bar{v}$ be optimal solutions for (1) and (6) respectively (By Theorem 2). Define $\bar{u}$ using (3) and (4) as $\bar{x} = B'\bar{u}$ and $\bar{u} = B\left(B'B\right)^{-1}\bar{x}$.

Now, $\bar{u}$ satisfies the constraints of (5) since,

$$AB'B\left(B'B\right)^{-1}\bar{x} = AB'BB^{-1}(B')^{-1}\bar{x} = A\bar{x} \geq b \quad \text{[Using (1)]}$$

Since $\bar{v}$ solves (6),

$$A'\bar{v} = c, \quad \bar{v} \geq 0 \tag{8}$$

Consequently,

$$BA'\bar{v} = Bc, \quad \bar{v} \geq 0 \tag{9}$$

Hence $\bar{v} \in W$, the dual feasible region of (7) and $\bar{u} \in U$. Consequently the dual pair (5) - (7) are both feasible and hence by Theorem 2 and Theorem 3 they are both solvable with equal extrema.

SHIV NADAR UNIVERSITY

Proof (Part 1)

Consequently,

$$c'B'\bar{u} = c'\bar{x} = \min_{x \in X} c'x = \max_{v \in V} b'v = b'\bar{v} \leq \max_{w \in W} b'w = \min_{u \in U} c'B'u \qquad (10)$$

The inequality follows from the fact that $\bar{v}$ is feasible for (7). The equality after that follows from the fact extrema of the primal - dual pair (5) - (7) are equal. So, we get

$$c'B'\bar{u} \leq \min_{u \in U} c'B'u$$

but since $\bar{u}$ is feasible for (5), it is indeed the optimal solution for (5).

Hence, $\bar{u} = B\left(B'B\right)^{-1}\bar{x}$ is the optimal solution for (5).

Proof (Part 2)

($\Leftarrow$)
Conversely, let $\bar{u}$ and $\bar{w}$ be optimal solutions for (5) and (7) respectively (By Theorem 2). Let $\bar{x} = B'\bar{u}$. Now, $\bar{x}$ satisfies (1) since $A\bar{x} = AB'\bar{u} \geq b$.
Since $\bar{w}$ solves (7) we have,

$$BA'\bar{w} = Bc, \bar{w} \geq 0 \tag{11}$$

Since, rank of B is n i.e the no.of columns, we can use the left inverse and it follows that,

$$A'\bar{w} = c, \bar{w} \geq 0 \tag{12}$$

Hence, $\bar{w} \in V$, the dual feasible region of (6) and $\bar{X} \in X$. Consequently the dual pair (1) - (6) are both feasible and hence by Theorem 2 and Theorem 3 they are both solvable with equal extrema.

Proof (Part 3)

We have shown that linear program (1) is solvable if and only if the secure linear program (5) is solvable. It remains to show that the extrema of these two linear programs are equal.

Since $\bar{w} \in W$ implies $\bar{w} \in V$, it follows that

$$\max_{w \in W} b'w = b'\bar{w} \leq \max_{v \in V} b'v \tag{13}$$

Hence,

$$\min_{u \in U} c'B'u = \max_{w \in W} b'w \geq \min_{x \in X} c'x = \max_{v \in V} b'v \geq \max_{w \in W} b'w = \min_{u \in U} c'B'u \tag{14}$$

where the equalities above follow from the equality of optimal primal and dual objectives of linear programs, the first inequality follows from (10) and the second inequality from (13).

Thus, the extremas of (1) and (5) are equal i.e

$$\min_{x \in X} c'x = \min_{u \in U} c'B'u.$$

The Algorithm

- All p entities agree on a $k \geq n$, the number of rows of the random matrix $B \in \mathbb{R}^{k \times n}$ as defined earlier.
- Each entity generates its own privately held random matrix $B_j \in \mathbb{R}^{k \times n_j}$, j = 1,..,p , where n_j is the number of features held by entity j which results in the matrix $B = [B_{.1} B_{.2} \ldots\ldots B_{.p}] \in \mathbb{R}^{k \times n}$
- Each entity j makes public only its matrix product $A_j B_j$ as well as its cost coefficient product $B_j c_j$. These products don't reveal either $A_{.j}$ or c_j but allow the public computation of the matrix and cost coefficient needed for the secure LP:

$$AB' = A_{.1}B'_{.1} + A_{.2}B'_{.2} + ... + A_{.p}B'_{.p}$$

$$c'B' = c'_1 B'_{.1} + c'_2 B'_{.2} + ... + c'_p B'_{.p}$$

The Algorithm

- A public optimal solution vector u to the secure LP (5) and a public optimal objective function value $c'Bu$ are computed. By the proposition , this optimal value equals the optimal objective function value of the original LP (1).
- Each entity computes its optimal solution vector x_j using $x = B'u$ as :

$$x_j = B'_{.j}u, j = 1, ..., p$$

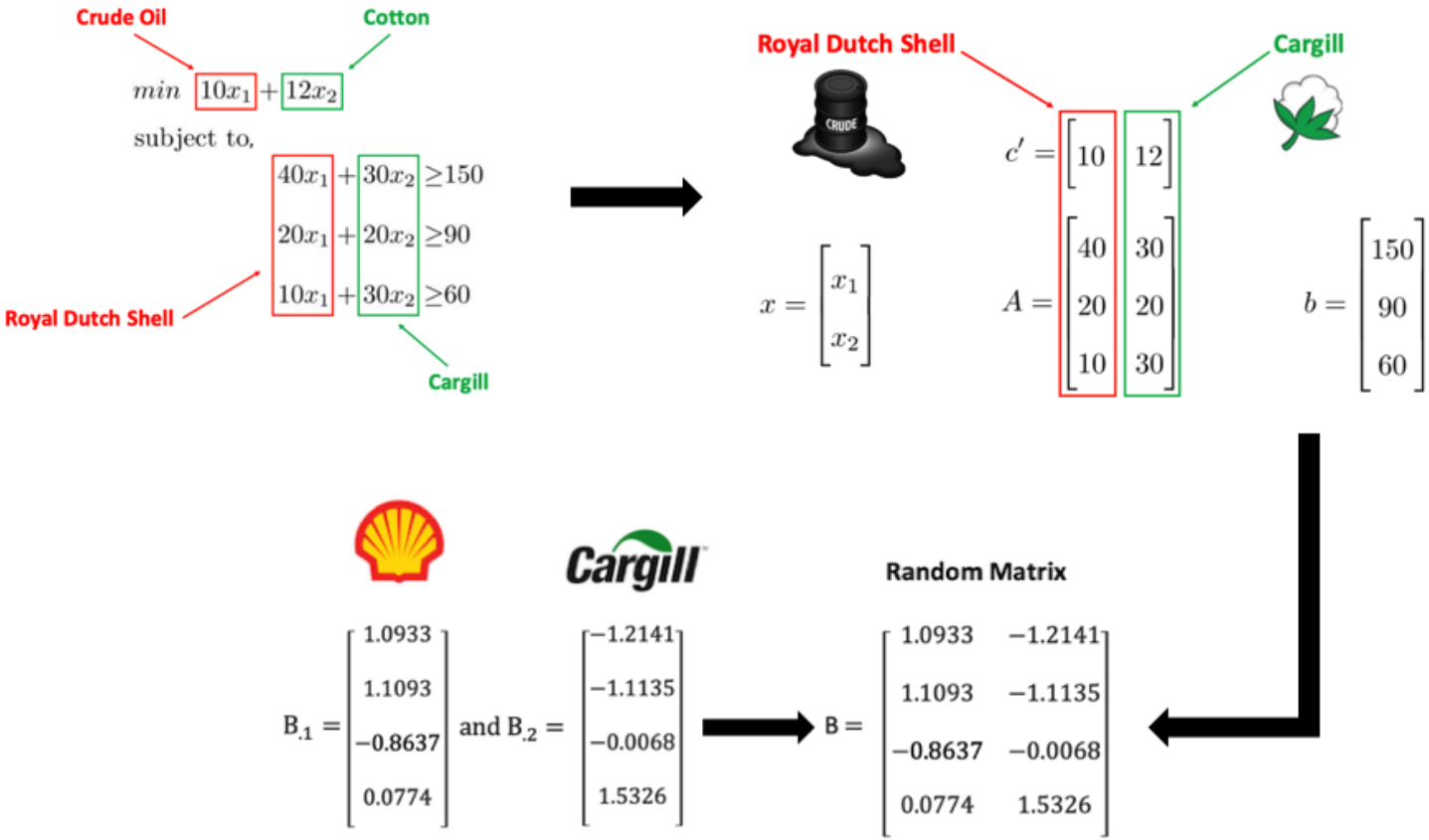
- The solution component vectors x_j, j = 1,...p, are revealed by its owners if a public solution vector to the original LP is agreed upon. Else the component vectors may be kept private if only the minimum value of (1) is needed, in which case that minimum value equals the the publicly available minimum value $\min_{u \in U} c'B'u$ of the secure LP (5).

Example 1

- Generate a random solvable LP with m = 100 and n = 1000.
- Partition columns of A and c into three groups with $n_1 = 500$, $n_2 = 300$ and $n_3 = 200$.
- Generate $B_{.1} \in \mathbb{R}^{n \times n_1}$, $B_{.2} \in \mathbb{R}^{n \times n_2}$, $B_{.3} \in \mathbb{R}^{n \times n_3}$

On solving the secure LP (5) and comparing its optimal objective value with that of (1), the two optimal objectives agreed to 14 significant figures. Computation time was 0.163 seconds. This was done using the CPLEX solver in MATLAB.

Example 2


Crude Oil
Cotton
min $10x_1 + 12x_2$
subject to,
$40x_1 + 30x_2 \geq 150$
$20x_1 + 20x_2 \geq 90$
$10x_1 + 30x_2 \geq 60$
Royal Dutch Shell
Cargill
Royal Dutch Shell
Cargill
$c' = [10 \quad 12]$
$x = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$
$A = \begin{bmatrix} 40 & 30 \\ 20 & 20 \\ 10 & 30 \end{bmatrix}$
$b = \begin{bmatrix} 150 \\ 90 \\ 60 \end{bmatrix}$
Random Matrix
$B_{.1} = \begin{bmatrix} 1.0933 \\ 1.1093 \\ -0.8637 \\ 0.0774 \end{bmatrix}$ and $B_{.2} = \begin{bmatrix} -1.2141 \\ -1.1135 \\ -0.0068 \\ 1.5326 \end{bmatrix}$
$B = \begin{bmatrix} 1.0933 & -1.2141 \\ 1.1093 & -1.1135 \\ -0.8637 & -0.0068 \\ 0.0774 & 1.5326 \end{bmatrix}$

Example 2

Solving the transformed LP we get,

$u_1 = 0, u_2 = 0.3635, u_3 = 0, u_4 = 3.5763$ and optimal value $= 46.5$

Using the transformation we get,

$x_1 = 3.75, x_2 = 0.75$ and optimal value $= 46.5$

References

1. Mangasarian, O.L. Privacy-preserving linear programming. Optim Lett 5, 165–172 (2011).
2. X. Feng and Z. Zhang. The rank of a random matrix. Applied Mathematics and Computation, 185:689–694, 2007.
3. Bazaraa MS, Jarvis JJ, Sherali HD (2011) Linear programming and Network Flows. Wiley, Hoboken.
4. https://en.wikipedia.org/wiki/Moore–Penroseinverse
5. https://www.mathworks.com/help/matlab/

SHIV NADAR UNIVERSITY