OM - Homework set-4

Akshay Bankar
2019201011

PAGE No.
DATE / /

Q-ID 102565

To solve $Ax = b$ given $A = \begin{bmatrix} 0 & 1 \\ 1 & 1 \\ 2 & 1 \end{bmatrix}$ and $b = \begin{bmatrix} 6 \\ 0 \\ 0 \end{bmatrix}$

(a) To check for consistency, we find rank of A and rank of augmented matrix [A|b]

$$[A|b] = \left[\begin{array}{cc|c} 0 & 1 & 6 \\ 1 & 1 & 0 \\ 2 & 1 & 0 \end{array}\right] \xrightarrow{R_3 \to R_3 - 2R_2} \left[\begin{array}{cc|c} 0 & 1 & 6 \\ 1 & 1 & 0 \\ 0 & 1 & 0 \end{array}\right]$$

$\downarrow R_3 \to R_3 - R_1$

$$\left[\begin{array}{cc|c} 0 & 1 & 6 \\ 1 & 1 & 0 \\ 0 & 0 & -6 \end{array}\right]$$

$\text{rank}(A) = 2 < \text{rank}(A|b) = 3$

Since $\text{rank}(A) < \text{rank}(A|b)$ the system of equations is inconsistent.

(b) Since the matrix $A_{m \times n}$ is such that $m > n$ i.e. no. of variables is less than number of equations, the system of equation is overdetermined. Hence we find least square solution which minimizes the square L2-norm $\|b - Ax\|_2^2$

Closed form solution of least square is given by

$$X = (A^T A)^{-1} A^T B.$$

$$A^T A = \begin{bmatrix} 0 & 1 & 2 \\ 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 1 & 1 \\ 1 & 2 \end{bmatrix} = \begin{bmatrix} 5 & 3 \\ 3 & 3 \end{bmatrix}$$

$$(A^T A)^{-1} = \frac{1}{6} \begin{bmatrix} 3 & -3 \\ -3 & 5 \end{bmatrix}$$

$$\therefore \; x = \frac{1}{6}\begin{bmatrix} 3 & -3 \\ -3 & 5 \end{bmatrix}\begin{bmatrix} 0 & 1 & 2 \\ 1 & 1 & 1 \end{bmatrix}\begin{bmatrix} 6 \\ 0 \\ 0 \end{bmatrix}$$

$$= \frac{1}{6}\begin{pmatrix} -3 & 0 & 3 \\ 5 & 2 & -1 \end{pmatrix}\begin{pmatrix} 6 \\ 0 \\ 0 \end{pmatrix}$$

$$= \frac{1}{6}\begin{pmatrix} -18 \\ 30 \end{pmatrix}$$

$$x = \begin{bmatrix} -3 \\ 5 \end{bmatrix}$$

answer

(c) Since the system is overdetermined, we optimize the least square i.e

$$\text{minimize} \; \| b - Ax \|_2^2$$

OM - Homework set-4

Akshay Bankar
2019201011

Q ID - 102566

To solve following system of linear equations -

$$\begin{bmatrix} 2 & -1 & 1 & -1 \\ 0 & 1 & -1 & -1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} 3 \\ 1 \end{bmatrix}$$

(i) No. of equations, $m = 2$
No. of variables, $n = 4$
Since $m < n$, i.e no. of equations are less than no. of variables, the system is under constrained.

(ii) Since the system is under constrained, multiple solutions exist.
~~The so~~ One of the solution is obtained by minimizing $\|x\|$. subject to $Ax = b$.
And the solution obtained by this has smallest norm of any solution.

$$\text{minimize } \|x\|$$
$$\text{s.t } \quad Ax = b, \quad \cancel{b=}$$

$$A = \begin{pmatrix} 2 & -1 & 1 & -1 \\ 0 & 1 & -1 & -1 \end{pmatrix} \quad b = \begin{pmatrix} 3 \\ 1 \end{pmatrix}$$

The closed form solution of this optimization problem is

$$x = A^T (AA^T)^{-1} b$$

~~(iii). $A^T (AA^T)$~~

$$AA^T = \begin{bmatrix} 2 & -1 & 1 & -1 \\ 0 & 1 & -1 & -1 \end{bmatrix} \begin{pmatrix} 2 & 0 \\ -1 & 1 \\ 1 & -1 \\ -1 & -1 \end{pmatrix} = \begin{bmatrix} 5 & -1 \\ -3 & 3 \end{bmatrix}, \quad (AA^T)^{-1} = \frac{1}{12} \begin{bmatrix} 3 & 1 \\ 3 & 5 \end{bmatrix}$$

$$A^T (AA^T)^{-1} = \begin{pmatrix} 2 & 0 \\ -1 & 1 \\ 1 & -1 \\ -1 & 1 \end{pmatrix} \times \frac{1}{12} \begin{bmatrix} 3 & 1 \\ 3 & 5 \end{bmatrix} = \frac{1}{12} \begin{pmatrix} 6 & 2 \\ 0 & 4 \\ 0 & -4 \\ -6 & -6 \end{pmatrix}$$

$$X = A^T (AA^T)^{-1} b = \begin{pmatrix} 6 & 2 \\ 0 & 4 \\ 0 & -4 \\ -6 & -6 \end{pmatrix} \begin{pmatrix} 3 \\ 1 \end{pmatrix} = \frac{1}{12} \begin{pmatrix} 20 \\ 4 \\ -4 \\ -12 \end{pmatrix}$$

$$\therefore \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} 5/3 \\ 1/3 \\ -1/3 \\ -1 \end{bmatrix}$$

(c) The quantity being minimized is the L2 norm of X.
i.e minimize $||X||$
s.t $Ax = b$

Hence ~~the~~ solution obtained is a least norm solution.

OM - Homework Set-4

Akshay Bankar
2019201011

PAGE No.
DATE / /

QID- 102567

(a) The given statement is true that the basic feasible solution obtained after a tie occurs in the pivot row choice during pivot step is degenerate. Following is the justification.

When solving using simplex method, at a step we have a set of basic variables x_B and set of non-basic x_s such that

$$z = \bar{z}_0 + \bar{c}_s x_s$$
$$x_B = \bar{b} - \bar{A}_s x_s$$

Here we choose x_s such that $c_s < 0$ and make z small by making x_s as large as possible.
But the value of x_s is bounded ~~by~~ so as to retain feasibility of $x_B \geq 0$.
Now an incoming variable x_s ~~is~~ ~~choosen~~ and the outgoing basic variable is chosen such that the ratio b_i/a_{is} is smallest i.e

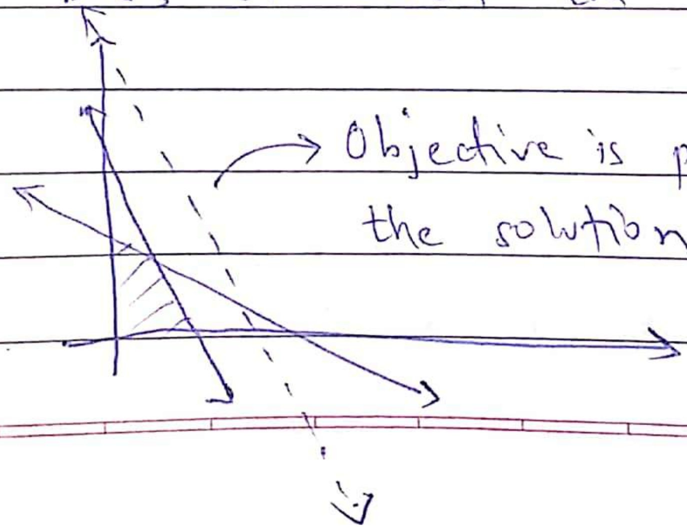
$$x_s = \frac{\bar{b}_r}{a_{rs}} = \min b_i/a_{is}$$

~~Now~~, if ~~$x_s = 0$~~, ~~one or k more~~ b_i ~~are~~ ~~zero~~ & the
Now if a tie occurs, ~~more~~ for choosing outgoing variable, more than one variables from set x_B will be zero and hence the feasible set obtained after this will have atleast one basic variable zero & hence ~~will~~ the solution will be degenerate.

(b) The given state is, in general, false. A different basic feasible solution is generated after every pivot step only in case where each step is non-degenerate. Though, when degenerate solutions occur, it is possible for $\bar{b}_r = 0$ (in the equation defined in (a)) in which case value of z does not decrease. And if the iterations are not discontinued, same selection of s and r will happen for each iteration & the same basic set would recur. And this recurrence can take place at every step & hence even the consecutive feasible solutions might not be distinct.

(c) The given statement that total number of optimal solutions is finite is false. Though the number of basis are finite, the ~~of~~ optimal solution may lie on one of the basis giving infinite optimal solutions if the objective is parallel to ~~the~~ one of the basis on which the ~~of~~ optimal solution lies.

Example:



OM Homework set-4

Akshay Bankar
2019201011

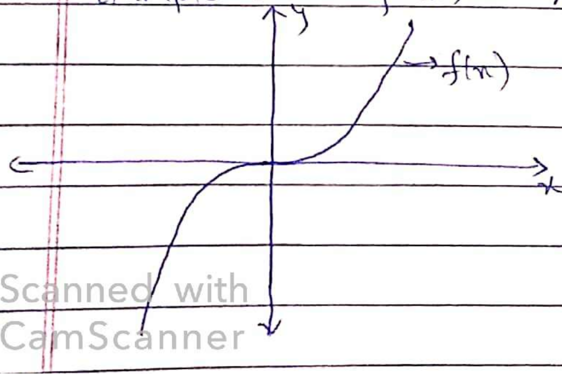
PAGE No.
DATE / /

ID - 102568

(i) a function that is both convex and concave

Assume function $f: \mathbb{R} \to \mathbb{R}$, $(x,y) \in \mathbb{R}^2$ and $\lambda \in [0,1]$
Then for a function to be convex,
$$f(\lambda x + (1-\lambda)y) \leq \lambda f(x) + (1-\lambda) f(y)$$
and for concavity,
$$f(\lambda x + (1-\lambda)y \geq \lambda f(x) + (1-\lambda) f(y)$$
Therefore, for a function to be both convex and concave,
$$f(\lambda x + (1-\lambda)y) = \lambda f(x) + (1-\lambda) f(y)$$
All linear functions follow this.
Ex:- $f(x) = x$
$$LHS = f(\lambda x + (1-\lambda)y) = \lambda x + (1-\lambda)y$$
$$RHS = \lambda f(x) + (1-\lambda)f(y) = \lambda x + (1-\lambda)y$$
Hence $LHS = RHS$, therefore $f(x) = x$ is both convex and concave.

(ii) a function that is neither convex nor concave.
Such a function can be created by stitching a concave function with a convex such that it is concave in one interval and convex in other.
Example :- $f(x) = x^3$



The function is concave on the left and convex on the right of y-axis.

OM - Homework Set-4

Akshay Bankar
2019201011

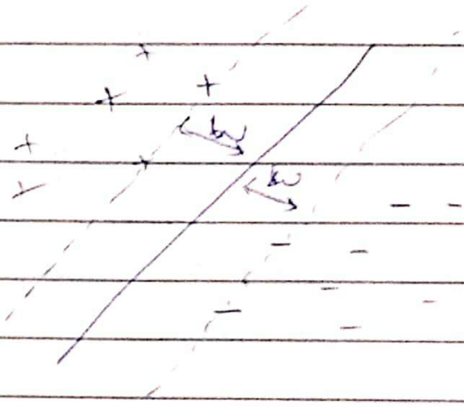
PAGE No.
DATE / /

Q-ID 102569

- To show that there exist infinite hyperplanes:

Given that there exists a hyperplane which separates two classes of points. This can be represented as

$wx + b < 0$ for one set of points
$wx + b > 0$ for other set of points



The above can be re-written as,

$wx + (b - m_1) = 0$
$wx + (b + m_2) = 0$
$m_1, m_2 \in \mathbb{R}$ & $m_1 m_2 > 0$

These two lines; when translated to both sides will still separate the two set of points. Now, m_1 & m_2 can take infinitly many real values. Hence there can be infinitely many hyperplanes obtained which separate these ~~to~~ points into two classes.

- Judging a good hyperplane:

A good hyperplane is the ~~one~~ that separates the points with as wide margin as possible. Therefore, the goal should be to ~~maximi~~ find such an hyperplane which maximizes the margin.

Primal formulation: To find the largest margin of separation, w which is given by $2/\|w\|$. Hence the formulation

$$\max_{w} \frac{2}{\|w\|} \quad \text{s.t} \quad \begin{array}{ll} w^T x_i + b \geq 1 & \text{if } y_i = +ve \text{ class} \\ w^T x_i + b \leq -1 & \text{if } y_i = -ve \text{ class} \end{array}$$

Or equivalently,

$$\min_{w} \|w\|^2 \quad \text{s.t} \quad y_i(w^T x_i + b) \geq 1 \quad \text{for } i = 1, \ldots, N$$

- When misclassification of some points is allowed i.e it is not forbidden to have points of one class on the opposite side of the hyperplane. In this case primal formulation changes as follows-

$$\min \|w\|^2 + C\frac{1}{n}\sum \varepsilon_i$$

$$\text{s.t} \quad y_i(w^T x_i + b) \geq 1 - \varepsilon_i$$

$$\forall_i \ \& \ \varepsilon_i \geq 0$$

~~where~~ The margin constraint is relaxed by introducing slack variables.

Akshay Bankar
2019201011

PAGE No.
DATE / /

OM - Homework set - 4

Q ID: 102570

KKT conditions: The Karush-Kuhn-Tucker conditions are first derivative tests for a solution in non-linear programming, to be optimal, provided that some regularity conditions are satisfied.

Necessary Conditions: Suppose that objective function ~~and~~ $f: \mathbb{R}^n \to \mathbb{R}$ and the constraint functions $g_i: \mathbb{R}^n \to \mathbb{R}$ and $h_j: \mathbb{R}^n \to \mathbb{R}$ are continuously differentiable at a pt. x^*. If x^* is local optima and the optimization problem satisfies some regularity conditions, then there exists constants μ_i $(i=1,\dots,m)$ & λ_j $(j=1,\dots,l)$, called KKT multipliers, such that following conditions hold:

(i) $f(x)$: $\nabla f(x^*) \mp \sum_{i=1}^{m} \mu_i \nabla g_i(x^*) - \sum_{j=1}^{l} \lambda_j \nabla h_j(x^*) = 0$

for maximizing (−) & minimizing (+).
This is stationarity.

(ii) Primal feasibility: $g_i(x^*) \leq 0$, for $i = 1,\dots,m$
$h_j(x^*) = 0$, for $j = 1,\dots,l$

(iii) Dual feasibility: $\mu_i \geq 0$ for $i = 1,\dots,m$

(iv) Complementary slackness:
$\mu_i g_i(x^*) = 0$ for $i = 1,\dots,m$

→ Cases when these conditions are sufficient for optimality:
The necessary ~~functions~~ conditions are sufficient for optimality if the objective function, ~~the inequality conf~~ f of a maximization problem is a concave function, the equality constraints g_j are continuously differentiable convex functions & the equality constraints h_i are affine functions. Similarly, if the objective function f of a minimization problem is convex function, the necessary conditions are sufficient for optimality.

OM - Homework Set-4

Akshay Bankar
2019201011

PAGE No.
DATE / /

QID - 102571

Given LP: $\min c^T x$

s.t $1^T x \le k$

$0 \le x_i \le 1, \quad i = 1, \dots, n$

The variables in this problem have upper bounds. Hence we solve this with simplex method for upper bound.
We find set of conditions which are useful in algorithm for bounded variables.
Complementary slackness condition - Let x^* be primal feasible & y^* be dual feasible. Then x^* is optimal and y^* is optimal iff.

$$x_j(C^T A_j - c_j) = 0 \quad \forall j$$

and

$$y_j(b_i - A_i x_i) = 0 \quad \forall i$$

Now x^* is primal optimal iff there exist dual feasible y^* such that x^* and y^* satisfy complementary slackness condition. Applying this for the given problem with upperbound,

$$\begin{array}{lll} \min c^T x & & \min c^T x \\ \text{s.t } 1^T x \le k & \Rightarrow & \text{s.t } 1^T x \le b \\ 0 \le x_i \le 1 & & x \le 1 \\ & & 0 \le x \end{array}$$

Let y and α be vectors of dual variables corresponding to $1^T x \le k$ and $x \le 1$, respectively. Then dual problem is,

$$\max K^T y + 1^T \alpha$$

$$\text{s.t } y^T + \alpha^T \ge C^T$$

$$y \ge 0, \quad \alpha \ge 0$$

Therefore by complementary slackness condition, a pair of feasible vectors x & $(y^T, \alpha^T)^T$ is optimal to primal & dual, respectively iff

$$x_j(y^T 1 + \alpha_j - c_j) = 0 \quad \forall j$$

$$y_i(k - 1^T x) = 0 \quad \forall i$$

and

$$\alpha_j(1 - x_j) = 0 \quad \forall j$$

OM - Homework - Set 4

Akshay Bankar
2019201011

PAGE No.
DATE / /

Q ID - 102572

Given problem: max $Z = 3x_1 + 4x_2$

s.t $2x_1 + x_2 \le 6$

$2x_1 + 3x_2 \le 9$

$x_1, x_2 \ge 0$

Convert it to canonical form by adding slack variables S_1 & S_2.

∴ the problem becomes max $Z = 3x_1 + 4x_2 + 0S_1 + 0S_2$

& s.t $2x_1 + x_2 + S_1 = 6$

$2x_1 + 3x_2 + S_2 = 9$

Iteration - 1

		C_j	3	4	0	0	
B	C_B	X_B	x_1	x_2	S_1	S_2	Min Ratio (X_B/x_2)
S_1	0	6	2	1	1	0	$6/1 = 6$
S_2	0	9	2	3	0	1	$9/3 = 3$
Z=0		Z_j	0	0	0	0	
		$Z_j - C_j$	-3	-4	0	0	

Negative minimum $Z_j - C_j$ is -4 which has column index 2. Hence, entering variable is x_2

Minimum ratio is 3 & its row index is 2. So, leaving variable is S_2.

∴ the pivot element is 3

Add $R_2(\text{new}) = R_2(\text{old})/3$

Add $R_1(\text{new}) = R_1(\text{old}) - R_2(\text{new})$

Iteration - 2

		C_j	3	4	0	0	
B	C_B	X_B	x_1	x_2	S_1	S_2	Min Ratio (X_B/x_1)
S_1	0	3	1.333	0	1	-0.333	$\frac{3}{1.333} = 2.25$
x_2	4	3	0.667	1	0	0.333	$\frac{3}{0.67} = 4.5$
Z=12		Z_j	2.667	4	0	1.333	
		$Z_j - C_j$	-0.333	0	0	1.333	

Negative min $Z_j - C_j$ is -0.333 & it which has column index 1. Hence the entering variable is x_1

Min ratio is 2.25 and its row index is 1. So the leaving variable is S_1

$\therefore$ The pivot element is 1.333

Add $R_1(new) = R_1(old) / 1.33$

Add $R_2(new) = R_2(old) - 0.667 \times R_1(new)$

Iteration - 3

		C_j	3	4	0	0
B	C_B	X_B	x_1	x_2	S_1	S_2
x_1	3	2.25	1	0	0.75	-0.25
x_2	4	1.5	0	1	-0.5	0.5
Z = 12.75		Z_j	3	4	0.25	1.25
		$Z_j - C_j$	0	0	0.25	1.25

Since all $Z_j - C_j \geq 0$, optimal solution is arrived with variable values as: $x_1 = 2.25$, $x_2 = 1.5$

and max $Z = 12.75$

OM - Homework set - 4

Akshay Bankar
2019201011

Q-ID 102573

Given that for a square matrix A, the row-sum of every row equals k.
To show: k is an eigen value.

- Let v be an eigenvector and λ be its corresponding eigenvalue ~~of~~ for the matrix A.

We need to find ~~such~~ a v such that,

$$Av = \lambda v \quad \& \quad \lambda = k$$

Now, consider a $v = [1 \; 1 \; 1 \; \dots \; 1]^T_{1 \times n}$

Hence,

$$Av = A_{n \times n} [1 \; 1 \; 1 \; \dots \; 1]^T$$

$$= \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ \vdots & & & \\ a_{n1} & \dots & & a_{nn} \end{pmatrix} \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix}_{n \times 1}$$

$$= \begin{pmatrix} a_{11} + a_{12} + \dots + a_{1n} \\ a_{21} + a_{22} + \dots \\ \vdots \\ a_{n1} + a_{n2} + \dots + a_{nn} \end{pmatrix}$$

It is given that $\sum_j a_{ij} \;\; \forall \; i = 1, \dots n$ is k

$$\therefore Av = k \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}_{n \times 1} = kv.$$

Hence k is an eigen value of A corresponding to the eigen vector $v = [1 \; 1 \; 1 \; \dots \; 1]^T_{1 \times n}$

OM - Homework Set - 4

Akshay Bankar
201920101

PAGE No.
DATE / /

Q ID - 102574

Given - A, B, C matrices such that,
$AB = BC$ and A and C have no common eigenvalues.

Sol^n - Suppose A and C matrix have a common eigenvalue λ.
Let u be the right eigenvector to A corresponding to eigenvalue λ

$\therefore \quad Au = \lambda u$

Let v be the left eigenvector to C corresponding to eigenvalue λ.

$\therefore \quad v^T C = \lambda v^T$

Now consider the following,

$(Au)v^T = (\lambda u)v^T = \lambda(uv^T)$ — ①

and

$u(v^T C) = u v^T \lambda = \lambda(uv^T)$ — ②

Given $AB = BC$ and comparing ① and ② we see that matrix B is uv^T.
Hence the matrix $B = uv^T$, But A & C have no eigenvalue in common, ~~therefore~~ therefore B must be zero. Hence B is a zero matrix.