John Min
jcm 2199

IEOR 6613

Optimization I

HW # 1

1.7, 1.11, 1.20, 2.5, 2.6, 2.10, 2.13, 2.14

$\boxed{1.7}$ **Moment problem :**

$z \in \{0, 1, \ldots, K\}$ w.p. $p_0, \ldots, p_K$

$\mathbb{E}[z] = \sum_{k=0}^{K} k \cdot p_k$, $\mathbb{E}[z^2] = \sum_{k=0}^{K} k^2 \cdot p_k$

$\mathbb{E}[z^4] = \sum_{k=0}^{K} k^4 \cdot p_k$

---

lower bound : minimize $\sum_{k=0}^{K} k^4 \cdot p_k$

s.t. $\sum_{k=0}^{K} p_k = 1$

$p_k \geq 0$ for $k \in \{0, 1, \ldots, K\}$ +

upper bound : minimize $-\sum_{k=0}^{K} k^4 p_k$

s.t. $\sum_{k=0}^{K} p_k = 1$

$p_k \geq 0$ for $k \in \{0, 1, \ldots, K\}$ +

add these 2 constraints to both UB, LB formulations

$\mathbb{E}[z] = \sum_{k=0}^{K} k p_k$

$\mathbb{E}[z^2] = \sum_{k=0}^{K} k^2 p_k$

$\boxed{1.11}$ **Optimal currency conversion :**

N currencies , exchange rate r_{ij} , limit u_i that can be exchanged , No arbitrage

B currency 1

maximize x_N $\iff$ min $-x_N$

s.t. $\sum_{j=2}^{N} c_{1j} \leq B$

$\sum_{i=1}^{N} c_{ij} \leq u_i$

$c_{ij} \geq 0$

$x_j = x_j + c_{ij} \cdot r_{ij}$ for $j = 1, \ldots, N$

$x_i = x_i - c_{ij} \cdot r_{ij}$ for $i = 1, \ldots, N$

$\boxed{1.20}$

(a) Let $S = \{Ax \mid x \in \mathbb{R}^n\}$ for given A.

S is a subspace of $\mathbb{R}^n$.

Pf:

DEF: A nonempty subset S of $\mathbb{R}^n$ is a subspace of $\mathbb{R}^n$ if $ax + by \in S$ $\forall x, y \in S$ and $\forall a, b \in \mathbb{R}$.

(i) By def., $\vec{0} \in \mathbb{R}^n$. Let $\vec{x} = \vec{0}$. Given any A, $A \cdot \vec{0} \in S$.

(ii) Again, trivially, for given matrix A, the vector $x_1 \in \mathbb{R}^n$, and scalar $c \in \mathbb{R}$, $c(Ax_1) \in S$. b/c $c \cdot Ax_1 = Ax_2$ where $x_2 = c \cdot x_1$. Since $Ax_2 \in S$ by definition, $c \cdot Ax_1 \in S$. Therefore, S is closed under scalar multiplication.

(iii) Let $a, b \in \mathbb{R}$. Choose $x, y \in \mathbb{R}^n$. Define $z = ax + by$. Clearly, $z \in \mathbb{R}^n$.
By def., $Az \in S$. Now, $Az = A(ax + by) = A \cdot ax + A \cdot by$. Hence, S is closed under addition + scalar multiplication.

By (i), (ii), (iii), $S \subseteq \mathbb{R}^n$.

(b) Assume S is a proper subspace of $\mathbb{R}^n$. Then, $\exists$ matrix B s.t. $S = \{y \in \mathbb{R}^n \mid By = 0\}$.

Pf:

DEF: S is a proper subspace $\iff$ S is a subspace of $\mathbb{R}^n$ and $S \neq \mathbb{R}^n$.

Since S is a proper subspace, $\exists$ non-zero vector $a \perp S$ s.t. $a'x = 0$ $\forall x \in S$.

Construct B s.t. its rows are composed of vectors that are orthogonal to vectors in $\mathbb{R}^n$.

Since $y \in \mathbb{R}^n$, we now have a matrix B s.t. our proper subspace of $\mathbb{R}^n$, S, can be defined as $S = \{y \in \mathbb{R}^n \mid By = 0\}$

(c) Suppose V is m-dimensional affine subspace of $\mathbb{R}^n$, $m < n$.
Show $\exists$ linearly independent vectors $a_1, \dots, a_{n-m}$ and scalars $b_1, \dots, b_{n-m}$ s.t. $V = \{y \mid a_i' y = b_i,\ i = 1, \dots, n-m\}$

Pf: Since V is an affine subspace of $\mathbb{R}^n$, $V = V_0 + x^0 = \{x + x^0 \mid x \in V_0\}$ for some x^0 where V^0 is a subspace of $\mathbb{R}^n$.

V is m-dimensional meaning that $V_0 = V - x^0 \neq \mathbb{R}^n$. $\Rightarrow$ V_0 is a proper subspace of $\mathbb{R}^n$.

In fact, V_0 has dimension $m < n$ and by def. of a proper subspace, $\exists$ non-zero vector $a \perp V_0$ s.t. $a'x = 0$ $\forall x \in V_0$. Given m-dimensionality, $\exists$ $n-m$ linearly independent vectors orthogonal to V_0.

Therefore, $V_0 = \{x \mid a_i' x = 0,\ i = 1, \dots, n-m\}$

Define $y := x + x^0$ and $b = x^0$. where $b_i = x_i^0$ for $i = 1, \dots, n-m$.

Then, we have linearly independent vectors $a_1, \dots, a_{n-m}$ & scalars $b_1, \dots, b_{n-m}$ s.t. $V = \{y \mid a_i' y = b_i,\ i = 1, \dots, n-m\}$

2.5 Extreme points of isomorphic polyhedra

DEF: A mapping f is affine if it is of form $f(x) = Ax + b$, where A is a matrix and b is a vector.

Let $P \in \mathbb{R}^n$, $Q \in \mathbb{R}^m$ whee P, Q are polyhedra.

DEF: P and Q are isomorphic if $\exists$ affine mappings $f: P \mapsto Q$ and $g: Q \mapsto P$ s.t.

$g(f(x)) = x \quad \forall x \in P$ and $f(g(y)) = y \quad \forall y \in Q$.

(a) Let P, Q be isomorphic. Then, $\exists$ one-to-one correspondence b/w their extreme points.
In particular, if f & g are as above, x is an extreme pt. of P $\Leftrightarrow$ $f(x)$ is an extreme pt of Q.

DEF: A vector $x \in P$, a polyhedron, is an extreme pt. of P if we cannot find two vectors $y, z \in P$ $y, z \neq x$, and scalar $\lambda \in [0,1]$ s.t. $x = \lambda y + (1-\lambda) z$. (i.e., not a convex combination of y, z).

Pf:

By THM 2.1 (b), Since P is a polyhedron, P is a convex set.

By THM 2.3, given a non-empty polyhedron P and $x^* \in P$, x^* is a vertex $\Leftrightarrow$ x^* is an extreme pt.

($\Rightarrow$)

Assume x^* be extreme pt. of P. Define $y^* = f(x^*)$

By THM 2.3, x^* is a vertex of P. By Def. 2.7, since x^* is a vertex of P, $\exists$ vector c s.t. $c'x < c'x^* \; \forall x \in P, x \neq x^*$.

For any $y \in Q$, $y \neq y^*$, $f(g(y)) = y \neq y^* = f(x^*) \Rightarrow g(y) \neq g(y^*) = x^*$ for $g(y) \in P$. (using P, Q's isomorphic properties).

Let affine function $g(y) := By + d$ for $B \in \mathbb{R}^{n \times m}$ and $d \in \mathbb{R}^n$.

Then, $c'(By+d) < c'(By^*+d) \quad \forall y \in Q, y \neq y^*$. $\Rightarrow (B'c)'y < (B'c)'y^* \quad \forall y \in Q, y \neq y^*$.

Suppose y^* is NOT an extreme pt. of Q.

Then, $y^* = \lambda y_1 + (1-\lambda) y_2$ for some $y_1, y_2 \in Q$, $y_1, y_2 \neq y^*$ and $\lambda \in (0,1)$ i.e., y^* is a convex combination of two other pts. in Q.

$\Rightarrow \nexists (B'c)'y < (B'c)'y^* \; \forall y \in Q, y = y^*$ which is a contradiction.

b/c $(B'c)'y^* = \lambda (B'c)'y_1 + (1-\lambda)(B'c)'y_2 < (B'c)'y^*$.

Contradiction $\Rightarrow$ $y^* = f(x^*)$ is an extreme pt. of Q.

($\Leftarrow$)

Assume $y^* = f(x^*)$ is an extreme pt. of Q. $\Rightarrow$ y^* is a vertex of P $\Rightarrow$ $\exists c$ s.t. $c'y < c'y^* \; \forall y \in P, y \neq y^*$.

Let $x^* = g(y^*)$.

For any $x \in P$, $x \neq x^*$, $g(f(x)) = x \neq x^* = g(y^*) \Rightarrow f(x) \neq f(x^*) = y^*$ for $f(x) \in P$. (by isomorphism of P, Q)

Given affine function $f(x) := Ax + b$ for $A \in \mathbb{R}^{m \times n}$, $b \in \mathbb{R}^m$,

$c'(Ax+b) < c'(Ax+b) \; \forall x \in P, x \neq x^* \Rightarrow (A'c)'x < (A'c)'x^* \; \forall x \in P, x \neq x^*$.

Similarly as above, if x^* is NOT an extreme pt., this above statement would not hold and we would have a contradiction, hence proving the ($\Leftarrow$) side of the argument. (the only if side)

$\boxed{2.5\ (b)}$ Introducing slack variables to isomorphic polyhedron

Set $P = \{x \in \mathbb{R}^n \mid Ax \geq b,\ x \geq 0\}$, $A \in \mathbb{R}^{k \times n}$

$Q = \{(x, z) \in \mathbb{R}^{n+k} \mid Ax - z = b,\ x \geq 0,\ z \geq 0\}$

P, Q are isomorphic.

Pf: Let f, g be affine mappings for $P \mapsto Q$, $Q \mapsto P$, respectively.

$f(x) := (x, Ax+b) \quad \forall x \in P$

$g(x, z) := x \quad \forall (x, z) \in Q.$

We can see the following:

$f(x) \in Q.$

$g(f(x)) = x \quad \forall x \in P.$

AND $g(x, z) \in P.$

$f(g(x, z)) = (x, z) \quad \forall (x, z) \in Q.$

Thus, by def., P, Q are isomorphic.

2.6 Carathéodry's Theorem

$A_1, \ldots, A_n$ are a collection of vectors in $\mathbb{R}^m$.

(a) Let $C = \left\{ \sum_{i=1}^{n} \lambda_i A_i \mid \lambda_1, \ldots, \lambda_n \geq 0 \right\}$

Show any element of C can be expressed in the form $\sum_{i=1}^{n} \lambda_i A_i$ w/ $\lambda_i \geq 0$ and w/ at most m of the coefficients λ_i being non-zero.

Pf:

Consider:

$\Lambda = \left\{ (\lambda_1, \ldots, \lambda_n) \in \mathbb{R}^n \mid \sum_{i=1}^{n} \lambda_i A_i = y, \ \lambda_1, \ldots, \lambda_n \geq 0 \right\}$ which for arbitrary $y \in \mathbb{R}^m$ is an element of C.

The polyhedron Λ is formulated as follows: $\underbrace{\begin{bmatrix} | & | & & | \\ A_1 & A_2 & \cdots & A_n \\ | & | & & | \end{bmatrix}}_{m \times n} \underbrace{\begin{bmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{bmatrix}}_{n \times 1} = y \in \mathbb{R}^m$

Let $A^T = \begin{bmatrix} A_1 & A_2 & \cdots & A_n \end{bmatrix}$, and $\lambda = \begin{bmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{bmatrix}$

$\Rightarrow A^T \lambda = y$

Assuming full row rank of A, ($\text{rank}(A) = m$), $\lambda = (A^T)^{-1} y$.

where $A^{-1} = \begin{bmatrix} A_B & A_N \end{bmatrix}$ where A_B has m non-zero vectors and A_N has $n-m$ zero vectors for their respective columns.

Hence, λ_i has m non-zero coefficients.

In the case where $\text{rank}(A) < m$, λ_i has $< m$ non-zero coefficients b/c

$A^{-1} = \begin{bmatrix} A_B & A_N \end{bmatrix}$ where A_B has $\text{rank}(A)$ non-zero vectors as columns

and A_N has $n - \text{rank}(A)$ zero columns.

Therefore, for $\lambda_i \geq 0$, λ_i has $\leq m$ non-zero coefficients.

$\boxed{2.6(b)}$ Carathéodry's THM :

Let P be the convex hull of vectors A_i : $P = \left\{ \sum_{i=1}^{n} \lambda_i A_i \;\middle|\; \sum_{i=1}^{n} \lambda_i = 1,\ \lambda_1, \dots, \lambda_n \geq 0 \right\}$

Then, any element of P can be expressed in the form $\sum_{i=1}^{n} \lambda_i A_i$ where $\sum_{i=1}^{n} \lambda_i = 1$, $\lambda_i \geq 0$ $\forall i$, with at most $m+1$ coefficients λ_i being non-zero

Pf: Analogous to part (a), consider an element of P :

Let $x \in P$, a point in the convex hull P.

Then, x is a convex combination of a finite number of points in P : $x = \sum_{i=1}^{n} \lambda_i x_i$, $\sum \lambda_i = 1$, $\lambda_i \in [0,1]$.

(i) Trivial case : $n \leq m+1$. ✓

(ii) Suppose $n > m+1$.

Then, consider the linear dependence of $x_2 - x_1, \dots, x_n - x_1$.

Therefore, $\exists\, \mu_2, \dots, \mu_n \in \mathbb{R}$ s.t. $\sum_{j=2}^{n} \mu_j (x_j - x_1) = 0$

Define $\mu_1 := -\sum_{j=2}^{n} \mu_j$.

Then, $\sum_{j=1}^{n} \mu_j x_j = 0$.

$\sum_{j=1}^{n} \mu_j = 0$. where not all $\mu_j = 0$. $\Rightarrow$ $\exists\, \mu_j > 0$ for $j \in \{1, \dots, n\}$.

Thus,

$$x = \sum_{i=1}^{n} \lambda_i x_i - c \cdot \sum_{i=1}^{n} \mu_i x_i = \sum_{i=1}^{n} (\lambda_i - c\mu_i) x_i \quad \text{for some } c \in \mathbb{R}.$$

Set $c := \min_{1 \leq i \leq n} \{ \lambda_i / \mu_i : \mu_i > 0 \} = \frac{\lambda_i}{\mu_i}$ and the equality will hold.

For $c > 0$, for $1 \leq i \leq n$, $\lambda_i - c\mu_i \geq 0$. By defining c as above, $\lambda_i - c\mu_i = 0$

Therefore,

$$x = \sum_{i=1}^{n} (\lambda_i - c\mu_i) x_i. \quad \text{where every } \lambda_i - c\mu_i \geq 0, \quad \sum_{i=1}^{n} (\lambda_i - c\mu_i) = 1,$$

and $\lambda_i - c\mu_i = 0$.

This means that x is represented as a convex combination of at most $n-1$ points of P. Since $n > m+1$, this is equivalent to at most $m+1$ non-zero coefficients of λ_i.

$\boxed{2.10}$ $P = \{x \mid Ax = b,\ x \geq 0\}$

$A \in \mathbb{R}^{m \times n}$ w/ linearly independent rows.

rank$(A) = m$.

(a) <u>If $n = m+1$, P has at most 2 BFS.</u>

TRUE.

Pf: $m = n-1$. The polyhedron P lies in an affine subspace w/ $m = n-1$ linearly independent constraints

Every sol'n of $Ax = b$ takes the form : $x_1 + cx_2$ where $x_1 \in P$ and $c \in \mathbb{R}$, where x_2 is a non-zero vector. This means the set P is contained in a line. If $\exists$ 3 BFS, the middle extreme pt. would be a convex combination b/w the 2 extreme pts $\Rightarrow$ 3rd BFS not an extreme pt. Contradiction. Therefore, $\exists \leq 2$ BFS for $n = m+1$.

(b) <u>The set of all optimal solutions is unbounded.</u>

FALSE.

Pf: Consider minimizing c s.t. $x \geq 0$, for arbitrary $c \in \mathbb{R}$. The optimal value is fixed BUT, The optimal sol'n set $x \in [0, \infty)$ is unbounded.

(c) <u>At every optimal sol'n, no more than m variables can be positive.</u>

FALSE.

Pf. Consider $c = 0$ in standard form.
Now, we can have n positive variables (assuming $n > m$)

(d) <u>If there are several optimal sol'ns, then $\exists$ at least 2 optimal BFS.</u>

FALSE.

Pf: A counter example would be one in which there are many optimal solutions but only only one BFS/vertex/extreme pt. Consider :

$$\min x_1 \quad \text{s.t. } x_1 \geq 0,\ x_2 \geq 0$$

Only the origin is a BFS. But, we have uncountably many optimal Solutions w/ $x_1 = 0$.

(e) <u>If there is more than one optimal sol'n, then there are uncountably many optimal solutions.</u>

TRUE.

Pf: b/c any convex combination of the optimal solutions are still optimal.
A line for example has uncountably many pts on the segment b/w its 2 endpoints.

(f) <u>Consider minimizing $\max\{c'x, d'x\}$ over P. If this has optimal sol'n, the optimal is at an extreme pt of P.</u>

FALSE.

Pf: The optimal sol'n need NOT occur at an extreme pt.

Consider : $\min \{ \max(1 - x_1, x_1 - 1) \}$

s.t. $x_1 \geq 0$

The optimal sol'n is at $x_1 = \frac{1}{2}$. BUT, the only extreme pt. of P is at $x = 0$.

$\boxed{2.13}$ $P = \{ x \mid Ax = b,\ x \geq 0 \}$

$A \in \mathbb{R}^{m \times n}$ w/ linearly independent rows

All BFS are nondegenerate in P.

Let $x \in P$ s.t. x has exactly m pos. components.

(a) x is a BFS

Pf: A is a full row rank $\left(\text{rank}(A) = m \right)$ and x has exactly m positive components $(x \in \mathbb{R}^m)$

$\Rightarrow$ all equality constraints are active.

$\Rightarrow$ we have m linearly independent constraints active at x.

By DEF 2.9, x is a basic solution.

All BFS are nondegenerat in P $\Rightarrow$ basic sol'n $x \in \mathbb{R}^m$ does NOT have more than m of the constraints active at x.

Therefore, x is a basic sol'n that satisfies all constraints.

$\Rightarrow$ x is a B.F.S.

(b) If $\nexists$ nondegeneracy assumption, x is not BFS.

We showed above that without the nondegeneracy assumption, we can only prove that x is a basic solution, but not necessarily a BFS.

Suppose $x \in \mathbb{R}^P$, $x \in P$ is B.F.S.

By DEF 2.9, x satisfies all constraints. + all equality constraints active + m linearly independent constraints active at x.

We are given $A \in \mathbb{R}^{m \times n}$ which is linearly independent $\Rightarrow$ full rank $\Rightarrow$ rank $(A) = m$.

AND x has exactly m positive components.

$\nexists$ $(m+1)^{th}$ constraint active at x. $\Rightarrow$ x not degenerate.

$\boxed{2.14}$ P is a bounded polyhedron in $\mathbb{R}^n$.

$\vec{a} \in \mathbb{R}^n$.

$b \in \mathbb{R}$

Define $Q := \{x \in P \mid a'x = b\}$

Every extreme pt. of Q is an extreme pt. of P or a convex combination of 2 adjacent extreme pts. of P.

Pf: Given $a \in \mathbb{R}^n$, $b \in \mathbb{R}$.

x^* is sol'n of $a'x = b$. Depending on # non-zero components of a, x^* can be a pt, a line, a hyperplane.

Q is the set of x^* intersecting polyhedron P.

(i) Suppose a has n non-zero components. Then, x^* is a pt. in P. since $\nexists$ convex combination b/w extreme pts in P, x^* is an extreme pt. in P. By DEF 2.9, x^* is a B.F.S. of Q by THM 2.6 since Q ~~hee~~ does NOT contain lines, just a set of extreme pts.

(ii) Consider the case that a has $< n$ non-zero components, call it $k < n$.

Then, x^* takes the form of a line or hyperplane.

Geometrically, $Q = x^* \wedge P$.

Since x^* now takes the form of a pt. and at least one vector.

In the direction of a's zero-components, sol'n x^* extends.

Yet, P is bounded and convex.

Thus, x^* will either hit an extreme pt. of P,

otherwise, it will hit an edge which is a c.c. of 2 adj. vertices of P.

This intersection is an extreme pt. of Q.