Massachusetts Institute of Technology
6.854/18.415: Advanced Algorithms
David Karger

Handout 17
Nov 11, 1999

Problem Set 8 Solutions

Problem 1. We are given polyhedra $P = \{x | Ax \leq b\}$ and $Q = \{x | Dx \leq d\}$.

(a) Among all pairs of points in P and Q, choose $x \in P, y \in Q$ that have the smallest distance $\|y - x\|$. We know that $\|y - x\| > 0$, since P and Q do not share common points. Let $c = y - x$.

Lemma 1 *Let $x \in P, y \in Q$ be such that $\|y - x\|$ is minimum. Then, any point $z \in P$ has $cz \leq cx$. Simiarly every $z \in Q$ has $cz \geq cy$.*

Proof. (By contradiction) Let $z \in P$ be such that $cz > cx$. The projection of $y - z$ along c is given by

$$\frac{c \cdot (y - z)}{c \cdot c} c = \frac{c \cdot (y - z)}{c \cdot (y - x)} (y - x) = k(y - x) \tag{1}$$

where $k > 1$. Let $y - z = k(y - x) + w$, where w is a vector orthogonal to c.

Since P is convex, the point $x' = (1 - \epsilon)x + \epsilon z$ for a small $\epsilon > 0$ will also lie in P. Now the distance from x' to y is

$$\begin{aligned} \|y - x'\| &= \|y - (1 - \epsilon)x - \epsilon z\| \\ &= \|(1 - \epsilon)(y - x) + \epsilon(y - z)\| \end{aligned}$$

Now, we can use the decomposition of z shown in Equation (1) to get

$$\begin{aligned} \|y - x'\| &= \|(1 - \epsilon)(y - x) + k\epsilon(y - x) + kw\| \\ &= \|(1 + \epsilon(k - 1))(y - x) + k\epsilon w\| \\ &= \sqrt{(1 + \epsilon(k - 1))^2 \|c\|^2 + k^2\epsilon^2 \|w\|^2} \end{aligned}$$

As $\epsilon \to 0$, the first order term of $\|y - x'\| - \|y - x\|$ is $\epsilon(k - 1) \|c\|$. Therefore x' is closer to y than x. This is a contradiction. So $cz \leq cx$ for all $z \in P$. Similarly, we can argue that $cz \geq cy$ for all $z \in Q$.

Corollary 1 *Let $x^* \in P, y^* \in Q$ be such that $\|y^* - x^*\|$ is minimum. Then, $c = y^* - x^*$ is such that $cx < cy$ for all $x \in P, y \in Q$.*

Proof. It is evident that $cx^* - cy^* = \|c\|^2 > 0$, since $x^* \neq y^*$. We can now apply Lemma 1 to the corollary.

(b) Constraints $Ax \leq b, Dx \leq d$ define the set $P \cap Q$. Therefore we can run an LP to find $\min\{0|Ax \leq b, Dx \leq d\}$ to find an $x \in P \cap Q$.

The solution to part (a) is not very useful in finding a separating hyperplane because the minimization constraint for the problem is not linear. But we know that a separating plane exists if $P \cap Q = \emptyset$. In other words, we know that there is a c such that $cx < cy$ for all $x \in P$ and $y \in Q$.

Consider the polyhedron

$$\begin{bmatrix} A & I & 0 \\ -D & 0 & -I \end{bmatrix} \begin{bmatrix} x \\ s \end{bmatrix} = \begin{bmatrix} b \\ d \end{bmatrix}$$

By Farkas' lemma, if the above equality is not satisfied, there exists a y such that

$$y \begin{bmatrix} A & I & 0 \\ -D & 0 & -I \end{bmatrix} \geq 0$$

$$y \begin{bmatrix} b \\ d \end{bmatrix} < 0$$

Let y be a solution to the above constraints. Any $Ax_1 \leq b$, $Dx_2 \leq d$ satisfies

$$y\,[A \quad -D] \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \leq yAx_1 - yDx_2 \leq yb - yd < 0$$

Therefore $y\,[A \quad -D]$ is a separating hyperplane. Computing such a y involves a single LP with objective 0.

Comments from graders: Part (b) was found to be hard. Solutions had to depend on Farkas' lemma since strong duality, although equivalent, was not introduced in this problem set. Many incorrect solutions to part (b) attempted to minimize the norm-2 distance using LP.

Problem 2.

(a) If x is optimal it is evident that no $x + \epsilon y$ can have a better solution. So $cy \geq 0$. We now prove the converse of this statement. If x is not optimal, then let z be the optimal point. Clearly all points of the form $(1 - \epsilon)x + \epsilon z$, for $\epsilon \in [0, 1]$ are in the polyhedron, since the polyhedron is convex. Thus we can consider a small $\epsilon \to 0^+$ so that $x + \epsilon(z - x)$ is feasible and $c(x+\epsilon(z-x)) \leq cx$. Therefore $z-x$ is a feasible direction with $c(z-x) \leq 0$.

(b) This proof is simlar to the proof for (a). An optimal x that is unique cannot have $cy < 0$ for a feasible direction y. Otherwise $c(x + \epsilon y) > cx$ and the optimality of x will be violated. The converse is proved by assuming a non-unique and possibly non-optimal solution x. Now a different optimal point z will be such that $cz \geq cx$, i.e., $c(z - x) \geq 0$. Now, let $y = z-x$ be the direction. Direction y is feasible since $x+\epsilon y$, for small $\epsilon \to 0^+$ is a convex combination of x and z. The value of $cy \leq 0$. We have thus shown that every non-unique and possibly non-optimal solution has a feasible direction that does not increase the objective function.

(c) Let us consider a feasible direction y. The point $x+\epsilon y$ will have a strictly better objective function iff $cy < 0$, from part (b). The objective function is given by $c_B A_B^{-1} B + \tilde{c}_N x_N$, where $\tilde{c}_N = c_N - c_B A_B^{-1} A_N$ gives the reduced costs. If all reduced costs are positive, the objective function will decrease iff at least one non-basic variable is decreased. But all non-basic variables are 0 and cannot be negative. The result follows.

Comments from graders: Most solutions to this problem were correct. Some failed to prove the existence of a feasible direction in parts (a) and (b).

Problem 3. Let n be the number of dimensions and m be the number of defining points in the polytope.

(a) A polytope is defined by the constraints

$$\begin{aligned} x &= \lambda \begin{bmatrix} x^{(1)} \\ \vdots \\ x^{(m)} \end{bmatrix} \\ \sum_{i=1}^{m} \lambda_i &= 1 \\ \lambda &\geq 0 \end{aligned}$$

where $x^{(1)}, \ldots, x^{(m)}$ are defining points of the polytope. Clearly the set of feasible points $(x_1, \ldots, x_n, \lambda_1, \ldots, \lambda_m)$ is a polyhedron. We still have to prove that the set of points $(x_1, \ldots, x_n)$ defines a polyhedron where $(x_1, \ldots, x_n)$ is a projection of some feasible point $(x_1, \ldots, x_n, \lambda_1, \ldots, \lambda_m)$. This can be proved by the repeated application of lemma 2 given below.

Lemma 2 *The projection of an $n+1$-dimensional polyhedron on n-dimensions is a polyhedron.*

Proof. Let us consider all constraints on the $(n+1)$st dimension x_{n+1}. The constraints, including equality constraints, can be written as

$$\begin{aligned} x_{n+1} &\geq f_l^{(i)}(x_1, \ldots, x_n, 1) \\ x_{n+1} &\leq f_g^{(j)}(x_1, \ldots, x_n, 1) \end{aligned} \tag{2}$$

where f_g and f_l define linear combinations on their parameters. Consider the constraints

$$f_l^{(i)}(x_1, \ldots, x_n, 1) \leq f_g^{(j)}(x_1, \ldots, x_n, 1) \qquad (\forall\, ij) \tag{3}$$

Any solution to constraints (2) gives a solution to contraints (3). Similarly any solution to (3) with

$$x_{n+1} \in [\max\{f_l^{(i)}(x_1, \ldots, x_n, 1) \mid i\}, \min\{f_g^{(j)}(x_1, \ldots, x_n, 1) \mid j\}]$$

which is a non-empty range, gives a feasible solution to (2).

(b) Let the point in the polyhedron be x. We will prove that x is linear combination of the vertices of the polyhedron by induction on the number of linearly independent constraints satisfied tightly. If n constraints are tight, the point x is the vertex defined by those constraints. Let the point x satisfy m tight constraints. Consider a line drawn through x and a vertex v satisfying those constraints. Such a vertex v exists, since any polyhedron with some constraints made tight is a polyhedron, if feasible. The existence of x proves feasibility. One point of intersection of this line with the boundary of the polyhedron is v. Let the other point of intersection of this line be x'. Now, x' belongs to at least one face that is not a linear combination of the m constraints that are already tight. Otherwise, the ray from x to x' can be extended infinitesimally without violating any constraint, thereby contradicting the definition of x'. Point x' now tightly satisfies $m + 1$ linearly independent constraints. The result follows from induction.

(c) The above inductive proof added one vertex to the convex sum per satisfied constraint. Initially, at least 0 constraints are tightly satisfied. The base case involved n constraints and expressed the point as a trivial convex combination of one vertex. Thus any point in the polyhedron can be expressed as a convex combination of $n + 1$ vertices.

Comments from graders: Many solutions to this problem, especially part (a), were not rigorous.

Problem 4. Since x is non-degenerate, it is defined by exactly n tight constraints. Every non-basic variable therefore defines a feasible direction, say $y^{(i)}$.

(a) Since x is the unique optimum, every feasible direction has $cy > 0$, from problem 2(b). So the feasible directions $y^{(i)}$ should have $\tilde{c}_N y^{(i)} > 0$, which is true only if every component in $\tilde{c}_N > 0$.

(b) Let z be the optimal vertex. Since x is a non-optimum vertex, $y = z - x$ is a feasible direction. The direction y has non-negative components for non-basic variables and can be expressed as a non-negative combination of $y^{(i)}$'s. The improvement in the objective function is given by $\tilde{c}_N y$ which is negative. Therefore at least one of the terms $\tilde{c}_N y^{(i)} < 0$. Thus some pivot step yeilds a strictly better solution.

Comments from graders: The solution to problem 2(b) cannot be directly applied in part (a), since some $y^{(i)}$ may not be feasible and could have non-positive reduced cost associated with it.

Problem 5. We write the problem as $Ax = b + e, x \geq 0$, where e is the vector of ϵ^i values.

(a) Since all rows are linearly independent, every set of basic variables B defines an invertible A_B. Let $B_1, \ldots, B_z$ denote all possible sets of basic variables. None of the vertices defined by these sets are equal if

$$A_{B_i}^{-1}(b + e) \neq A_{B_j}^{-1}(b + e) \quad (\forall\, ij)$$

Consider the equations defined by

$$A_{B_i}^{-1}(b + e) = A_{B_j}^{-1}(b + e) \quad (\forall\, ij)$$

Consider any equality defined above. Since $B_i \neq B_j$, there is atleast one power of ϵ that has a non-zero coefficient. Thus each equality defines some roots for ϵ. An ϵ smaller than the smallest positive root will ensure that all vertices defined by basic variables are unequal, thereby ensuring that there are no degenerate vertices.

(b) Let B^* be the basis of an optimum basic feasible solution to the perturbed problem. The optimal solution is unique if ϵ is sufficiently small (as defined by part (a)). The vector of reduced costs for the perturbed problem is given by $c_N - c_B A_B^{-1} A_N$, which is the same for the original problem. Since the solution to the perturbed problem is unique, all reduced costs are positive (from problem 4(a)). From problem 2(b), the solution to the original problem is optimal.

Comments from graders: Some solutions failed to account for degeneracy due to linearly dependent rows in A that cannot be corrected by perturbation.