10725/36725 Optimization
Recitation 1, Sep 4

Aaditya Ramdas
aramdas@cs.cmu.edu

1 Definitions

Vector algebra - Vector space (closed under addition, scalar multiplication, subspace), metric space (triangle inequality, positivity, symmetry?, zero iff equal?), normed space (positivity?, cauchy-schwarz, triangle ineq, defines a metric), inner product space (linearity, symmetry, positive definite?, defines a norm).

Linear algebra - positive definite matrices (equivalent definitions), cones, matrix dot product, vector norms (2,1,infinity) and matrix norms (frobenius, spectral, nuclear). SVD, PCA, EVD, column space, row space, null space, under and over-constrained linear systems.

Probability, Calculus - gradients, hessians, jacobians, expected values and linearity, taylor's theorem, law of iterated expectations, convergence (of sequences).

2 Alternate Simplex Algo

For $\max_x c^\top x$ subject to $Ax \leq b$.

Setup Note that maximimizing $c^\top x$ on $x^\top x \leq 1$ is achieved by going along direction c until we hit the edge of the set (or the hyperplane is tangent to the sphere). Similarly, maximizing it on a polygon, it is achieved at the furthest point in direction c. Generalizing this to higher dimensions, we can see that the optimum will be at the corner of $Ax \leq b$ where A is an $m \times n$ matrix and the set is the intersection of $m > n$ halfspaces in n dimensions (like a square in 2-D, or a cuboid in 3-D). Assume for simplicity of exposition that this set is full-dimensional (has non-zero n-dimensional volume), it is bounded, it has no degeneracies (no more than n hyperplanes pass through the same point in n dimensional space). Convince yourself that this is a convex set and that every point is either an extreme point or vertex of this closed and bounded polyhedron, or it can be written as a convex combination of extreme points (vertices), and that a linear function is maximized at some extreme point.

Method Assume we start at any feasible vertex. Such a vertex satisfies $A'x_0 = b'$ for some subset A', b' of n equations out of the m given by A, b. I claim that x_0 has n neighbours, each corresponding to swapping out one row of A' for some other row in $A'^c = A \setminus A'$. We are going to hop from neighbour to neighbour while we strictly increase the attained value.

Let us calculate the neighbours by walking along one of the n lines emanating from x_0 to its neighbours $x_1, ..., x_n$. The i-th line is given by $A'_i x < b'_i, A'_j x = b'_j$ if $j \neq i$ (if a vertex is the intersection of n hyperplanes, then a line is the intersection of $n-1$ of them). Let us call the directions of the lines $Z = [x_1 - x_0, x_2 - x_0, ..., x_n - x_0]$. Noting that $A'(x_1 - x_0)$ is going to look like $[< 0, 0, 0, 0..., 0]^\top$, we can find unknown directions Z by solving $A'Z = -I$. Hence, the directions towards the neighbours are given by the columns of $-(A')^{-1}$. Note that these are only directions towards the vertices and not the actual vertices.

Now, choose any direction such that $c^\top z_i > 0$ (any point on the line from x_0 in direction z_i is given by $x_0 + \lambda z_i$ for some $\lambda > 0$, and we would be ensuring that $c^\top(x_0 + \lambda z_i) > c^\top x_0$ which is the progress condition we want). How much can we move in this direction z_i? Till we make some other inequality constraint into an equality (if we moved beyond this, we would violate $Ax \leq b$). So we need the largest value of lambda that still satisfies $A(x_0 + \lambda z_i) \leq b$. This is given by $\lambda = \min_{j \in A'^c} (b - A_j x_0)/A_j z_i$.

What if there is no column with $c^\top z_i > 0$ (ie, there is no neighbour with higher cost or $c^\top x_i \leq x^\top x_0$)? Then we are at the global maximum. Why? Look at a small neighbourhood around x_0. Convince yourself that every point there can be written as a convex combination of x_0 and its neighbours i.e. $p = \sum_{i=0}^n \lambda_i x_i$ with $\sum_i \lambda_i = 1$. Hence $c^\top p = \sum_i \lambda_i (c^\top x_i) \leq c^\top x_0 \sum_i \lambda_i = c^\top x_0$ implying that p is the local maximum, and hence the global maximum. Can also show this by arguing that $-(A')^{-1}$ is full rank and hence forms a basis.