Homework 3

Sahit Mandala

February 17, 2016

Problem 1

Let some sequences of vectors $\{u_k\}, \{v_k\}$ be given. Suppose that some subsequence of $\{u_k\}$, say $\{w_k\} \subseteq \{u_k\}$, is bounded away from 0. Hence we can choose $\epsilon > 0$ and $K \in \mathbb{N}$ s.t. $||w_k - 0|| = ||w_k|| > \epsilon$ for all $k \geq K$. We know that $w_k = A_k v_k + o(1)$, so we can equivalently write $w_k = ||w_k||\hat{w}_k = A_k v_k + o(1)$ where $\hat{w}_k$ is merely the directional vector of norm 1. So $\hat{w}_k = \frac{A_k v_k}{||w_k||} + \frac{o(1)}{||w_k||} = A_k \frac{v_k}{||w_k||} + \frac{o(1)}{||w_k||}$. With some more manipulation, we see that $(\hat{w}_k - \frac{o(1)}{||w_k||}) = A_k \frac{v_k}{||w_k||}$. Notice $\frac{o(1)}{||w_k||} \leq \frac{o(1)}{\epsilon}$.

Recall that $||A * x|| \leq ||A|| * ||x||$ (where $||A|| = max_{||x||=1}||Ax||$, the matrix norm). Point is, we see that $||A_k \frac{v_k}{||w_k||}|| \leq ||A_k|| * ||\frac{v_k}{||w_k||}||$ and because $A_k \to A$, we know that $||A_k - A|| \to 0$ and $||A|| > 0$ since A is nonsingular. Setting $\alpha = \frac{3||A||}{2}$, we can choose a sufficiently large $K \in \mathbb{N}$ s.t. $||A_k - A|| \leq \frac{||A||}{2}$, which implies $||A_k|| \leq ||A_k - A|| + ||A|| \leq \frac{3||A||}{2} = \alpha > 0$. We also know that $\frac{v_k}{||w_k||} \to 0$, so we can choose a large enough $K^{'} \geq K$ s.t. $||\frac{v_k}{||w_k||}|| \leq \frac{1}{3\alpha}$ for $\alpha > 0$. Overall, for $k \geq K^{'}$, we see that $||A_k \frac{v_k}{||w_k||}|| \leq ||A_k|| * ||\frac{v_k}{||w_k||}|| \leq \alpha * \frac{1}{3\alpha} = \frac{1}{3}$.

We can also choose $K^{''} \geq K^{'} \in \mathbb{N}$ s.t. $||\frac{o(1)}{||w_k||}|| \leq ||\frac{o(1)}{\epsilon}|| \leq \frac{1}{3}$. By the triangle inequality, we know that $||\hat{w}_k|| \leq ||\hat{w}_k - \frac{o(1)}{||w_k||}|| + ||\frac{o(1)}{||w_k||}||$, so $||\hat{w}_k|| = 1 \leq ||\hat{w}_k - \frac{o(1)}{||w_k||}|| + ||\frac{o(1)}{||w_k||}|| \leq ||\hat{w}_k - \frac{o(1)}{||w_k||}|| + \frac{1}{3}$ thus and $1 - \frac{1}{3} = \frac{2}{3} \leq ||\hat{w}_k - \frac{o(1)}{||w_k||}||$. Be we can also see that this is a contradiction since $(\hat{w}_k - \frac{o(1)}{||w_k||}) = A_k \frac{v_k}{||w_k||}$ but their norms $\frac{2}{3} \leq ||\hat{w}_k - \frac{o(1)}{||w_k||}||$ and $||A_k \frac{v_k}{||w_k||}|| \leq \frac{1}{3}$ do not agree on $k \geq K'$. Hence, our assumption was false and there was no such subsequence which was bounded away from 0.

Problem 2

Suppose that we have x^* which satisfies the MFCQ conditions at a point.

Consider the following dual problem:

max $\sum_{i \in A \cap I} \lambda_i$ s.t.

$\sum_{i \in E} \nabla c_i(x)\lambda_i + \sum_{i \in A \cap I} \nabla c_i(x)\lambda_i = \nabla f(x^*)$, $\lambda_i \geq 0$ for $i \in A \cap I$

This dual problem encapulates the constraints on the Lagrange multipliers as per KKT conditions given the $c_i(x)$. The primal for this problem would be (since $\lambda_i, i \in E$ are free):

min $\nabla f(x^*)^T d$ s.t.

$\nabla c_i(x)^T d_i = 0$ for $i \in E$

$\nabla c_i(x)^T d_i \geq 1$ for $i \in A \cap I$

d free

We know my MFCQ that we can choose some $d^{'}$ s.t. $\nabla c_i(x)^T d_i^{'} = 0$ for $i \in E$ and $\nabla c_i(x)^T d_i^{'} > 0$ for $i \in A \cap I$. So consider the point $\overline{d} = \frac{d^{'}}{min_{i \in A \cap I}(\nabla c_i(x)^T d_i^{'})} \geq 0$. Then $\nabla c_i(x)^T \overline{d}_i = \frac{\nabla c_i(x)^T d_i^{'}}{min_{i \in A \cap I}(\nabla c_i(x)^T d_i^{'})} = \frac{0}{min_{i \in A \cap I}(\nabla c_i(x)^T d_i^{'})} = 0$ for $i \in E$ and $\nabla c_i(x)^T \overline{d}_i = \frac{\nabla c_i(x)^T d^{'}}{min_{i \in A \cap I}(\nabla c_i(x)^T d_i^{'})} \geq \frac{min_{i \in A \cap I}(\nabla c_i(x)^T d_i^{'})}{min_{i \in A \cap I}(\nabla c_i(x)^T d_i^{'})} = 1$ for $i \in A \cap I$. So the primal problem is feasible at $\overline{d}$. Hence, this implies that the dual problem is bounded above. Recall that $\lambda_i \geq 0$ for $i \in A \cap I$ in the dual problem, so because $\sum_{i \in A \cap I} \lambda_i$ is bounded above for all feasible points (which are just all valid Lagrange multipliers), it readily follows that each λ_i is bounded above for $i \in A \cap I$ (each are nonnegative and their sum is bounded above, so none of them can diverge to infinite without causing the sum to become unbounded).

With this statement, since the $\overline{\lambda_i}, i \in E$ are bounded, the linear linear combinations of $\nabla c_i(x), i \in E$ with these coefficients must also be bounded (say by some vector V s.t. $||V|| \geq \sum_{i \in A \cap I} \nabla c_i(x) \lambda_i$ for all Lagrange Multipliers) . So $\sum_{i \in E} \nabla c_i(x) \lambda_i = \nabla f(x^*) - \sum_{i \in A \cap I} \nabla c_i(x) \lambda_i$ and notably $\nabla f(x^*) - \sum_{i \in A \cap I} \nabla c_i(x) \lambda_i$ is always bounded. This implies that the vector $\sum_{i \in E} \nabla c_i(x) \lambda_i$ is also bounded for every Lagrange Multiplier. Because $\nabla c_i(x)$ are linearly independent by MFCQ, it readily follows that the $\lambda_i, i \in E$ are bounded.

Problem 3

Consider the problem as given:

min $f(x)$ s.t. $g(x) \geq 0, h(x) \geq 0, g(x)^T h(x) = 0$

For $g(x) \geq 0$, we know that each entry $g(x)_i \geq 0$, and for $h(x) \geq 0$, we know that each entry $h(x)_i \geq 0$. We can use these expanded constraints on real valued function $g_i(x), h_i(x)$. Hence, the new but equivalent problem would be:

min $f(x)$ s.t. $g_i(x) \geq 0, h_i(x) \geq 0, g(x)^T h(x) = 0$ for all $i \in \{1, ..., m\}$

Let some feasible point x^* be given. So $g_i(x^*) \geq 0, h_i(x^*) \geq 0, g(x^*)^T h(x^*) = 0$ for all $i \in \{1, ..., m\}$. But note that $g(x^*)^T h(x^*) = \sum_{i \in \{1,...,m\}} g_i(x) * h_i(x) = 0$. And because, for each i, $g_i(x^*) \geq 0, h_i(x^*) \geq 0$, it follows that $g_i(x^*) * h_i(x^*) \geq 0$. So for $\sum_{i \in m} g_i(x) * h_i(x) = 0$, clearly $g_i(x^*) * h_i(x^*) = 0$ for each $i \in \{1, ..., m\}$. Since each term is nonnegative, it follows that $g_i(x^*) = 0$ or $h_i(x^*) = 0$ for each $i \in \{1, ..., m\}$. Let S be the set of active constraint gradients. For each $i \in \{1, ..., m\}$, we know that either $\nabla g_i(x^*) \in S$ if $g_i(x^*) = 0$ or $\nabla h_i(x) \in S$ if $h_i(x^*) = 0$ and we know atleast one of those clauses will always be true. Hence, we have atleast m active constraint gradients in S. Furthermore we also know that the constraint $c(x) = g(x)^T h(x) = 0$ is also active, so $\nabla c(x) \in S$. But because there are $m + 1$ active constraint gradients in S in an $n \leq m$ dimensional space, we know that these vectors must be linearly dependent (assumes $n \leq m$).

The last gradient $\nabla g(x^*)^T h(x^*) = \nabla \sum_{i \in \{1,...,m\}} g_i(x^*) h_i(x^*) = \sum_{i \in \{1,...,m\}} \nabla (g_i(x^*) h_i(x^*)) = \sum_{i \in \{1,...,m\}} \nabla g_i(x^*) h_i(x^*) + g_i(x^*) \nabla h_i(x^*) \in S$. Since either $h_i(x^*) = 0$ or $g_i(x^*) = 0$, we know that either $\nabla g_i(x^*) h_i(x^*) + g_i(x^*) \nabla h_i(x^*) = g_i(x^*) \nabla h_i(x^*)$ or $\nabla g_i(x^*) h_i(x^*) + g_i(x^*) \nabla h_i(x^*) = \nabla g_i(x^*) h_i(x^*)$ respectively (for any i). If $h_i(x^*) = 0$, then $\nabla h_i(x^*) \in S$, so $\nabla h_i(x^*)$ and $\nabla g_i(x^*) h_i(x^*) + g_i(x^*) \nabla h_i(x^*)$ are linearly dependent (with factor $g(x^*)$. Similarly if $g(x^*) = 0$. So each term can be written as a linear combination of vectors in S, and thus the sum $\nabla g(x^*)^T h(x^*) = \sum_{i \in \{1,...,m\}} \nabla g_i(x^*) h_i(x^*) + g_i(x^*) \nabla h_i(x^*)$ is linearly dependent on the other vectors in S. Hence, we see (in a more general case) that S is linearly dependent.

Problem 4

A)

In this case, LICQ does not hold. Notice that because x^* satisfies $c_i(x^*) = 0$, for each i, it readily follows that $-c_i(x^*) = -1 * 0 = 0$. So in the reformulated problem, both constraints are active at x^*. But the then the set of active constraint gradients are $\{\nabla c_i(x^*), -\nabla c_i(x^*)\}_{i \in \{1,...,m\}}$, and clearly $\nabla c_1(x^*) + (-\nabla c_1(x^*)) = 0$, showing that these gradients are in fact linearly dependent at x^*.

B)

From the original problem, we know that $c_i(x^*) = 0, \forall i \in \{1, ..., m\}$ and $\lambda_i \geq 0, \forall i \in \{1, ..., m\}$ since these were the Lagrange multipliers for the original problem's KKT conditions at x^*.

In the revised problems, there are no equality constraints, only inequalities. Define $c_i'(x) = c_i(x), \forall i \in \{1, ..., m\}$ and $c_{i+m}'(x) = c_i(x), \forall i \in \{1, ..., m\}$. Notice that $c_i(x^*) = -c_i(x^*) = 0, \forall i \in \{1, ..., m\}$ from above, so clearly $c_i'(x^*) \geq 0, \forall i \in \{1, ..., 2m\}$. For each $i \in \{1, ..., m\}$, we can choose the Lagrange multipliers as follows: if $\lambda_i' \geq 0$, then $\lambda_i' = \lambda_i, \forall i \in \{1, ..., m\}$

and $\lambda_{i+m}^{'} = 0, \forall i \in \{1, ..., m\}$; otherwise, $\lambda_i^{'} = 0, \forall i \in \{1, ..., m\}$ and $\lambda_{i+m}^{'} = \lambda_i, \forall i \in \{1, ..., m\}$. By this choice, $\lambda_i^{'} \geq 0$ for all i. Furthermore, we see that $\lambda_i^{'} c_i^{'}(x) = \lambda_i^{'} c_i(x^*) = \lambda_i^{'} * 0 = 0$ for $i \in \{1, .., m\}$ and $\lambda_{i+m}^{'} c_{i+m}^{'}(x) = \lambda_i^{'} * -c_i(x^*) = \lambda_i^{'} * 0 = 0$ for all $i \in \{1, ..., m\}$. Furthermore, $\sum_{i \in A} \nabla c_i^{'}(x) \lambda_i^{'} = \sum_{i \in A} \nabla c_i(x) \lambda_i = \nabla f(x^*)$ from the original problem due to our alternating choice of $\lambda_i^{'}$ s.t. a positive λ_i is matched with $c_i(x^*)$ while a negative λ_i is converted into $-\lambda_i$ and matched with a $-c_i(x^*)$. This construction satisfies the KKT conditions and forms a valid Lagrange multiplier.

However, the Lagrange Multipliers are not unique in the reformulated problem. Because the constraint gradients are linearly dependent as we showed above, there are multiple potential LMs. For example, we could. For each $i \in \{1, ..., m\}$, we can choose the Lagrange multipliers as follows: if $\lambda_i^{'} \geq 0$, then $\lambda_i^{'} = \lambda_i + 1, \forall i \in \{1, ..., m\}$ and $\lambda_{i+m}^{'} = 1, \forall i \in \{1, ..., m\}$; otherwise, $\lambda_i^{'} = 01, \forall i \in \{1, ..., m\}$ and $\lambda_{i+m}^{'} = \lambda_i + 1, \forall i \in \{1, ..., m\}$. Then we see that, termwise, either $\nabla c_i^{'}(x) \lambda_i^{'} + \nabla c_{i+m}^{'}(x) \lambda_{i+m}^{'} = \nabla c_i(x)(1) + -(\lambda_i + 1) - \nabla c_i(x) = \nabla c_i(x)(\lambda_i)$ or $\nabla c_i^{'}(x) \lambda_i^{'} + \nabla c_{i+m}^{'}(x) \lambda_{i+m}^{'} = \nabla c_i(x)(\lambda_i + 1) + -\nabla c_i(x) = \nabla c_i(x)(\lambda_i)$, which amounts to the same sum as above.

C)

In general, no. So we know that for some w, Suppose some w exists which satisfies the MFCQ. That would imply that $\nabla c_i^{'}(x^*)^T w > 0$ for all $i \in \{1, ..., 2m\}$. So consider $\nabla c_1^{'}(x^*) = \nabla c_1(x^*)$ and $\nabla c_{m+1}^{'}(x^*) = -c_1(x^*)$. So $\nabla c_1(x^*)^T w > 0$ and $(-\nabla c_1(x^*))^T w > 0$. But this implies that $(-\nabla c_1(x^*))^T w = -(\nabla c_1(x^*)^T w) > 0$ and $\nabla c_1(x^*)^T w > 0$ simultaneously, a contradiction