MATH 407 Key Theorems

Theorem 0.1 (Weak Duality Theorem)**.** *If $x \in \mathbb{R}^n$ is feasible for $\mathcal{P}$ and $y \in \mathbb{R}^m$ is feasible for $\mathcal{D}$, then*

$$c^T x \le y^T Ax \le b^T y.$$

Thus, if $\mathcal{P}$ is unbounded, then $\mathcal{D}$ is necessarily infeasible, and if $\mathcal{D}$ is unbounded, then $\mathcal{P}$ is necessarily infeasible. Moreover, if $c^T\bar{x} = b^T\bar{y}$ with $\bar{x}$ feasible for $\mathcal{P}$ and $\bar{y}$ feasible for $\mathcal{D}$, then $\bar{x}$ must solve $\mathcal{P}$ and $\bar{y}$ must solve $\mathcal{D}$.

Proof. Let $x \in \mathbb{R}^n$ be feasible for $\mathcal{P}$ and $y \in \mathbb{R}^m$ be feasible for $\mathcal{D}$. Then

$$\begin{array}{rcll}
c^T x & = & \sum\limits_{j=1}^{n} c_j x_j & \\
& \le & \sum\limits_{j=1}^{n} (\sum\limits_{i=1}^{m} a_{ij} y_i) x_j & [\text{since } 0 \le x_j \text{ and } c_j \le \sum\limits_{i=1}^{m} a_{ij} y_i, \text{ so } c_j x_j \le (\sum\limits_{i=1}^{m} a_{ij} y_i) x_j] \\
& = & y^T A x & \\
& = & \sum\limits_{i=1}^{m} (\sum\limits_{j=1}^{n} a_{ij} x_j) y_i & \\
& \le & \sum\limits_{i=1}^{m} b_i y_i & [\text{since } 0 \le y_i \text{ and } \sum\limits_{j=1}^{n} a_{ij} x_j \le b_i, \text{ so } (\sum\limits_{j=1}^{n} a_{ij} x_j) y_i \le b_i y_i] \\
& = & b^T y &
\end{array}$$

To see that $c^T\bar{x} = b^T\bar{y}$ plus $\mathcal{P}$–$\mathcal{D}$ feasibility implies optimality, simply observe that for every other $\mathcal{P}$–$\mathcal{D}$ feasible pair (x, y) we have

$$c^T x \le b^T\bar{y} = c^T\bar{x} \le b^T y\ .$$

□

Theorem 0.2. [The Fundamental Theorem of Linear Programming] Every LP has the following three properties:

(i) If it has no optimal solution, then it is either infeasible or unbounded.
(ii) If it has a feasible solution, then it has a basic feasible solution.
(iii) If it is bounded, then it has an optimal basic feasible solution.

Proof. **Part (i):** Suppose an LP has no solution. This LP is either feasible or infeasible. Let us suppose it is feasible. In this case, the first phase of the two-phase simplex algorithm produces a basic feasible solution. Hence, the second phase of the two-phase simplex algorithm either discovers that the problem is unbounded or produces an optimal basic feasible solution. By assumption, the LP has no solution so it must be unbounded. Therefore, the LP is either infeasible or unbounded.

Part (ii): If an LP has a feasible solution, then the first phase of the two-phase simplex algorithm produces a basic feasible solution.

Part (iii): Suppose an LP is bounded. In particular, this implies that the LP is feasible, and, so by Part (ii), it has a basic feasible solution. The second phase of the two-phase simplex algorithm either discovers that the problem is unbounded or produces an optimal basic feasible solution. Since the LP is bounded, the second phase produces an optimal basic feasible solution. □

Theorem 0.3 (The Strong Duality Theorem). *If either $\mathcal{P}$ or $\mathcal{D}$ has a finite optimal value, then so does the other, the optimal values coincide, and optimal solutions to both $\mathcal{P}$ and $\mathcal{D}$ exist.*

Proof. Since the dual of the dual is the primal, we may as well assume that the primal has a finite optimal value. In this case, the Fundamental Theorem of Linear Programming says that an optimal basic feasible solution exists. By our formula for the general form of simplex tableaus, we know that there exists a nonsingular record matrix $R \in \mathbb{R}^{n \times n}$ and a vector $y \in \mathbb{R}^m$ such that the optimal tableau has the form

$$\begin{bmatrix} R & 0 \\ -y^T & 1 \end{bmatrix} \begin{bmatrix} A & I & b \\ c^T & 0 & 0 \end{bmatrix} = \begin{bmatrix} RA & R & Rb \\ c^T - y^T A & -y^T & -y^T b \end{bmatrix}.$$

Since this is an optimal tableau, we know that

$$c - A^T y \le 0, \qquad\qquad -y^T \le 0$$

with $y^T b$ equal to optimal value in the primal problem. But then $A^T y \ge c$ and $0 \le y$ so that y is feasible for the dual problem $\mathcal{D}$. In addition, the Weak Duality Theorem implies that

$$\begin{array}{rcll} b^T y & = & \text{maximize} & c^T x & \le \; b^T \widehat{y} \\ & & \text{subject to} & Ax \le b, \; 0 \le x & \end{array}$$

for every vector $\widehat{y}$ that is feasible for $\mathcal{D}$. Therefore, y solves $\mathcal{D}$. Therefore, optimal solutions to both $\mathcal{P}$ and $\mathcal{D}$ exist with the optimal values coinciding. □

Theorem 0.4 (Fundamental Theorem on Sensitivity Analysis). *If $\mathcal{P}$ is primal nondegenerate, i.e. the optimal value is finite and no basic variable in any optimal tableau takes the value zero, then the dual solution y^* is unique and there is an $\epsilon > 0$ such that*

$$V(u) = b^T y^* + u^T y^* \quad \textit{whenever} \;\; |u_i| \le \epsilon, \; i = 1, \ldots, m \; .$$

Thus, in particular, the optimal value function V is differentiable at $u = 0$ with $\nabla V(0) = y^*$.

Proof. Let

$$\begin{bmatrix} RA & R & Rb \\ (c - A^T y^*)^T & -(y^*)^T & -b^T y^* \end{bmatrix}$$

be any optimal tableau for $\mathcal{P}$. Primal nondegeneracy implies that every component of the vector Rb is strictly positive. If there is another dual optimal solution $\tilde{y}$ associated with another tableau, then we can pivot to it using simplex pivots. All of these simplex pivots must be degenerate since the optimal value cannot change. But degenerate pivots can only be performed if the tableau is degenerate, i.e. there is an index i such that $(Rb)_i = 0$. But then the basic variable associated with $(Rb)_i$ must take the value zero contradicting the hypothesis that Rb is a strictly positive vector. Hence the only possible optimal tableau is the one given. The only other way to have multiple dual solutions is if there is an unbounded ray of optimal solutions emanating from the optimal solution identified by the unique optimal tableau. For this to occur, there must be a row in the optimal tableau such that any positive multiple of that row can be added to the objective row without changing the optimal value. Again, this can only occur if some $(Rb)_i$ is zero leading to the same contradiction. Therefore, primal nondegeneracy implies the uniqueness of the dual solution y^*.

Next let $0 < \delta < \min\{(Rb)_i \,|\, i = 1, \ldots, m\}$. Due to the continuity of the mapping $u \to Ru$, there is an $\epsilon > 0$ such that $|(Ru)_i| \le \delta \; i = 1, \ldots m$ whenever $|u_j| \le \epsilon \; j = 1, \ldots, n$. Hence, if we perturb b by u, then

$$R(b + u) = Rb + Ru \ge Rb - \epsilon \mathsf{e} > 0$$

whenever $|u_j| \le \epsilon \; j = 1, \ldots, n$, where e is the vector of all ones. Therefore, if we perturb b by u in the optimal tableau with $|u_j| \le \epsilon \; j = 1, \ldots, n$, we get the tableau

$$\begin{bmatrix} RA & R & Rb + Ru \\ (c - A^T y^*)^T & -(y^*)^T & -b^T y^* - u^T y^* \end{bmatrix}$$

which is still both primal and dual feasible, hence optimal with optimal value $V(u) = b^T y^* + u^T y^*$ proving the theorem. □

Theorem 0.5. *[Fundamental Theorem of the Alternative] Given $A \in \mathbb{R}^{m\times n}$, we have*

$$\mathrm{Ran}(A) = \mathrm{Nul}(A^T)^\perp, \quad \mathrm{Nul}(A) = \mathrm{Ran}(A^T)^\perp$$
$$\mathrm{Ran}(A^T) = \mathrm{Nul}(A)^\perp, \text{ and } \mathrm{Nul}(A^T) = \mathrm{Ran}(A)^\perp.$$

Proof. Observe that

$$\begin{array}{rcl}\mathrm{Nul}(A) & := & \{x \mid Ax = 0\} \\ & = & \{x \mid A_{i\cdot} \bullet x = 0,\ i = 1, 2, \ldots, m\} \\ & = & \{A_{1\cdot}, A_{2\cdot}, \ldots, A_{m\cdot}\}^\perp \\ & = & \mathrm{Span}\,[A_{1\cdot}, A_{2\cdot}, \ldots, A_{m\cdot}]^\perp \\ & = & \mathrm{Ran}(A^T)^\perp\ .\end{array}$$

Hence, $\mathrm{Nul}(A) = \mathrm{Ran}(A^T)^\perp$, and since for any subspace $S \subset \mathbb{R}^n$, we have $(S^\perp)^\perp = S$, we also have $\mathrm{Nul}(A)^\perp = \mathrm{Ran}(A^T)$. By replacing A by A^T, we obtain the remaining equivalences in the statement of the theorem. □

Theorem 0.6. *[Existence and Uniqueness for the Linear Least Squares Problem]*
Let $A \in \mathbb{R}^{m\times n}$ and $b \in \mathbb{R}^m$ and consider the linear least squares problem

$$\mathcal{LLS} \qquad \min_{x\in\mathbb{R}^n} \frac{1}{2}\|Ax - b\|_2^2\ .$$

1. *The vector $\bar{x}$ solves $\mathcal{LLS}$ if and only if $A^TAx = A^Tb$.*
2. *A solution to the normal equations $A^TAx = A^Tb$ always exists.*
3. *A solution to the linear least squares problem $\mathcal{LLS}$ always exists.*
4. *The linear least squares problem $\mathcal{LLS}$ has a unique solution if and only if $\mathrm{Nul}(A) = \{0\}$ in which case $(A^TA)^{-1}$ exists and the unique solution is given by $\bar{x} = (A^TA)^{-1}A^Tb$.*

Proof. 1. Suppose that $\bar{x}$ is a solution to $\mathcal{LLS}$, i.e.,

$$\|A\bar{x} - b\|_2 \le \|Ax - b\|_2 \quad \forall\, x \in \mathbb{R}^n. \tag{1}$$

Let x be any other vector in $\mathbb{R}^n$. Then

$$\begin{aligned}\|A\bar{x} - b\|_2^2 &= \|A(\bar{x} - x) + (Ax - b)\|_2^2 \\ &= \|A(\bar{x} - x)\|_2^2 + 2(A(\bar{x} - x))^T(Ax - b) + \|Ax - b\|_2^2 \\ &\ge \|A(\bar{x} - x)\|_2^2 + 2(A(\bar{x} - x))^T(Ax - b) + \|A\bar{x} - b\|_2^2 \qquad \text{(by (1))}.\end{aligned} \tag{2}$$

Therefore, by canceling $\|A\bar{x} - b\|_2^2$ from both sides, we know that, for all $x \in \mathbb{R}^n$,

$$0 \ge \|A(\bar{x} - x)\|_2^2 + 2(A(\bar{x} - x))^T(Ax - b) = 2(A(\bar{x} - x))^T(A\bar{x} - b) - \|A(\bar{x} - x)\|_2^2.$$

By setting $x = \bar{x} - tw$ for $t \in T$ and $w \in \mathbb{R}^n$, we find that

$$\frac{t^2}{2}\|Aw\|_2^2 \ge tw^TA^T(A\bar{x} - b) \quad \forall\, t \in \mathbb{R} and w \in \mathbb{R}^n.$$

Dividing by $t > 0$, we find that

$$\frac{t}{2}\|Aw\|_2^2 \ge w^TA^T(A\bar{x} - b) \quad \forall\, t > 0 and w \in \mathbb{R}^n,$$

and sending t down to zero gives

$$0 \ge w^TA^T(A\bar{x} - b) \quad \forall\, w \in \mathbb{R}^n,$$

which implies that $A^T(A\bar{x} - b) = 0$, or equivalently,

$$A^TA\bar{x} = A^Tb. \tag{3}$$

Next assume that $A^TA\bar{x} = A^Tb$. Then, for all $x \in \mathbb{R}^n$,

$$\begin{aligned}
\|Ax - b\|_2^2 &= \|(Ax - A\bar{x}) + (A\bar{x} - b)\|_2^2 \\
&= \|A(x - \bar{x})\|_2^2 + 2(A(x - \bar{x}))^T(A\bar{x} - b) + \|A\bar{x} - b\|_2^2 \\
&\geq 2(x - \bar{x})^T A^T(A\bar{x} - b) + \|A\bar{x} - b\|_2^2 \qquad (\text{since } \|A(x - \bar{x})\|_2^2 \geq 0) \\
&= \|A\bar{x} - b\|_2^2 \qquad (\text{since } A^T(A\bar{x} - b) = 0),
\end{aligned}$$

or equivalently, $\bar{x}$ solves $\mathcal{LLS}$.

2. Since $\mathrm{Ran}(A^TA) = \mathrm{Ran}(A^T)$, a solution to $A^TAx = A^Tb$ must exist.

3. By part (1), $\bar{x}$ solves $\mathcal{LLS}$ if and only if $A^TAx = A^Tb$. Hence the result follows from part (2).

4. The $\bar{x}$ solves the linear least squares problem if and only if $\bar{x}$ solves the normal equations. Hence, the linear least squares problem has a uniques solution if and only if the normal equations have a unique solution. Since $A^TA \in \mathbb{R}^{n\times n}$ is a square matrix, this is equivalent to saying that A^TA is invertible, or equivalently, $\mathrm{Nul}(A^TA) = \{0\}$. However, $\mathrm{Nul}(A) = \mathrm{Nul}(A^TA)$. Therefore, the linear least squares problem has a uniques solution if and only if $\mathrm{Nul}(A) = \{0\}$ in which case A^TA is invertible and the unique solution is given by $\bar{x} = (A^TA)^{-1}A^Tb$. $\square$