AM 221: Advanced Optimization	Spring 2016
Prof. Yaron Singer	*Section 4 — Wednesday, Feb. 17th*

1 Sensitivity Analysis

Let us consider the following linear program in canonical form:

$$\begin{aligned} \min_{\mathbf{x}\in\mathbb{R}^n} \ & \mathbf{c}^\intercal \mathbf{x} \\ \text{s.t } & A\mathbf{x} = \mathbf{b} \\ & \mathbf{x} \geq 0 \end{aligned}$$

for $A \in \mathbb{R}^{m\times n}$, $\mathbf{c} \in \mathbb{R}^n$ and $\mathbf{b} \in \mathbb{R}^m$. Sensitivity analysis concerns itself with the following general question:

To which extent are optimal solutions and optimal values
of optimization problems sensitive to changes in their parameters?

This question is important because in many scenarios, the parameters of an optimization problem come from the "environment" and are either varying or uncertain. Quantifying the robustness of a solution to changes of the parameters can save the cost of having to re-solve the optimization problem to adapt to those changes.

More precisely, for the above problem, let us consider an optimal basic feasible solution $\mathbf{x}^*$. We saw in class that we can assume without loss of generality that $\mathbf{x}^{*\intercal} = [\mathbf{x}_B^\intercal \ 0]$, where $A = [B \ N]$ with $B \in \mathbb{R}^{m\times m}$ invertible. Recall that $\mathbf{x}^*$ is:

1. Feasible iff $\mathbf{x}_B = B^{-1}\mathbf{b} \geq 0$.
2. Optimal iff $\mathbf{r}_N = \mathbf{c}_N^\intercal - \mathbf{c}_B^\intercal B^{-1} N \geq 0$.

We will consider perturbations of the constraint vector $\mathbf{b}$ and the cost vector $\mathbf{c}$ and give conditions under which $\mathbf{x}^*$ stays feasible and optimal for the new perturbed problem. A similar analysis could be conducted for changes in the matrix A.

1.1 Perturbation of the constraint vector b

In this section, we consider a perturbation where the vector $\mathbf{b}$ is replaced by $\mathbf{b}(\lambda) \stackrel{\text{def}}{=} \mathbf{b} + \lambda\mathbf{b}'$ for some perturbation direction $\mathbf{b}' \in \mathbb{R}^m$ and scaling factor $\lambda \in \mathbb{R}$. To follow the perturbation, $\mathbf{x}_B$ needs to be replaced with $\mathbf{x}(\lambda) \stackrel{\text{def}}{=} B^{-1}(\mathbf{b} + \lambda\mathbf{b}') = B^{-1}\mathbf{b} + \lambda B^{-1}\mathbf{b}'$.

Note that the optimality criterion 2. above is unaffected by the perturbation since it does not depend on $\mathbf{b}$. As a consequence if $\mathbf{x}(\lambda)$ is feasible for the new perturbed problem, it will be optimal.

$\mathbf{x}(\lambda)$ is feasible if $B^{-1}\mathbf{b} + \lambda B^{-1}\mathbf{b}' \geq 0$, that is:

$$\forall\, 1 \leq i \leq m,\ \lambda(B^{-1}\mathbf{b}')_i \geq -(B^{-1}\mathbf{b})_i$$

This can be rewritten as:

$$\begin{cases} \lambda \geq -\frac{(B^{-1}\mathbf{b})_i}{(B^{-1}\mathbf{b}')_i} & \text{if } (B^{-1}\mathbf{b}')_i > 0 \\ \\ \lambda \leq -\frac{(B^{-1}\mathbf{b})_i}{(B^{-1}\mathbf{b}')_i} & \text{otherwise} \end{cases}$$

This discussion can be summarized in the following proposition:

Proposition 1. *For any $\lambda \in \mathbb{R}$ and $\mathbf{b}' \in \mathbb{R}^m$, define:*

$$\underline{\lambda} \overset{\text{def}}{=} \max\left\{ -\frac{(B^{-1}\mathbf{b})_i}{(B^{-1}\mathbf{b}')_i} \,\middle|\, 1 \leq i \leq m \wedge (B^{-1}\mathbf{b}')_i > 0 \right\}$$

$$\bar{\lambda} \overset{\text{def}}{=} \min\left\{ -\frac{(B^{-1}\mathbf{b})_i}{(B^{-1}\mathbf{b}')_i} \,\middle|\, 1 \leq i \leq m \wedge (B^{-1}\mathbf{b}')_i < 0 \right\}$$

Then if $\lambda \in [\underline{\lambda}, \bar{\lambda}]$, solution $\mathbf{x}(\lambda) \overset{\text{def}}{=} \mathbf{x}_B + \lambda B^{-1}\mathbf{b}$ is feasible and optimal for the perturbed problem where $\mathbf{b}$ is replaced by $\mathbf{b} + \lambda \mathbf{b}'$.

Remark 2. It is also interesting to look at how the optimal value of the problem is affected by the perturbation. We have:

$$\mathbf{c}_B^\intercal \mathbf{x}(\lambda) = \mathbf{c}_B^\intercal \mathbf{x}_B + \lambda \mathbf{c}_B^\intercal B^{-1}\mathbf{b}$$

In other words, the new optimal value is obtained from the original optimal value by adding a linear function of λ. The rate of this linear function, namely $\mathbf{c}_B^\intercal B^{-1}\mathbf{b}$, has a natural interpretation using the dual of the unperturbed problem. Let $\mathbf{w}$ be an optimal solution of the dual, then using the Complementary Slackness conditions (cf. Problem Set 4), we have $\mathbf{w}^\intercal = \mathbf{c}_B^\intercal B^{-1}$. The rate can then simply be expressed as $\mathbf{w}^\intercal \mathbf{b}$.

1.2 Perturbation of the cost vector c

We will now consider the case where the cost vector $\mathbf{c}$ is replaced with $\mathbf{c} + \lambda \mathbf{c}'$ for some perturbation direction $\mathbf{c}' \in \mathbb{R}^n$ and some scaling factor λ. We note that the feasibility condition $\mathbf{x}_B = B^{-1}\mathbf{b} \geq 0$ is unaffected by this change, hence we will not need to modify our solution $\mathbf{x}_B$ to adapt to this perturbation. However, modifying $\mathbf{c}$ might break the optimality criterion. The solution will be optimal for the new perturbed problem as long as:

$$\left(\mathbf{c}_N^\intercal + \lambda \mathbf{c}_N^{'\intercal}\right) - \left(\mathbf{c}_B^\intercal + \lambda \mathbf{c}_B^{'\intercal}\right) B^{-1}N \geq 0$$

Reordering the terms, this can be re-written as:

$$\left(\mathbf{c}_N^\intercal - \mathbf{c}_B^\intercal B^{-1}N\right) + \lambda\left(\mathbf{c}_N^{'\intercal} - \mathbf{c}_B^{'\intercal} B^{-1}N\right) = \mathbf{r}_N + \lambda \mathbf{r}'_N \geq 0$$

where we defined $\mathbf{r}'_N \stackrel{\text{def}}{=} \mathbf{c}'^{\intercal}_N - \mathbf{c}'^{\intercal}_B B^{-1} N$. Similarly to perturbations of $\mathbf{b}$, this leads to the following conditions:

$$\begin{cases} \lambda \geq -\frac{(\mathbf{r}_N)_i}{(\mathbf{r}'_N)_i} & \text{if } (\mathbf{r}'_N)_i > 0 \\ \lambda \leq -\frac{(\mathbf{r}_N)_i}{(\mathbf{r}'_N)_i} & \text{otherwise} \end{cases}$$

and we can summarize this discussion in the following proposition.

Proposition 3. *For any $\lambda \in \mathbb{R}$ and $\mathbf{c}' \in \mathbb{R}^m$, define:*

$$\underline{\lambda} \stackrel{\text{def}}{=} \max \left\{ -\frac{(\mathbf{r}_N)_i}{(\mathbf{r}'_N)_i} \,\Big|\, m+1 \leq i \leq n \wedge (\mathbf{r}'_N)_i > 0 \right\}$$

$$\bar{\lambda} \stackrel{\text{def}}{=} \min \left\{ -\frac{(\mathbf{r}_N)_i}{(\mathbf{r}'_N)_i} \,\Big|\, m+1 \leq i \leq n \wedge (\mathbf{r}'_N)_i < 0 \right\}$$

Then if $\lambda \in [\underline{\lambda}, \bar{\lambda}]$, solution $\mathbf{x}_B$ is feasible optimal for the perturbed problem where $\mathbf{c}$ is replaced by $\mathbf{c} + \lambda \mathbf{c}'$.

Remark 4. Similarly to the Section 1.1, we can look at the optimal value of the new perturbed problem:

$$(\mathbf{c}_B + \lambda \mathbf{c}'_B)^{\intercal} \mathbf{x}_B = \mathbf{c}_B^{\intercal} \mathbf{x}_B + \lambda \mathbf{c}'^{\intercal}_B B^{-1} \mathbf{b}$$

Again, the new value is obtained from the original value by adding a linear function of λ. However, the rate of this function no longer has a natural interpretation in terms of a dual optimal solution.

2 Review of multivariate calculus

Topics covered: gradient, hessian matrix, Taylor expansion, first and second order characterization of local optima. Please refer to the lecture notes of Feb 22nd for more details.