Barrier functions, Duality, Sensitivity

Computational Intelligence, Lecture 13

by Sergei Savin

Spring 2022

Content

- Barrier functions
- Analytic center of linear inequalities
- Lagrange dual function
- Duality gap, strong and weak duality
- Sensitivity
- Homework

Linear inequalities

Consider linear inequality constraints:

$$\mathbf{A}\mathbf{x} \leq \mathbf{b} \tag{1}$$

Remember that we can rewrite it as:

$$\mathbf{a}_i^\top \mathbf{x} \leq b_i \tag{2}$$

$$\mathbf{a}_i^\top \mathbf{x} - b_i \leq 0 \tag{3}$$

Instead of *hard constraints* in (3) we can turn these into a cost function component:

$$J = -\sum_{i=1}^{n} \log(b_i - \mathbf{a}_i^\top \mathbf{x}) \tag{4}$$

Which is called a *barrier function*.

Let us consider barrier functions $J = -\sum_{i=1}^{n} \log(b_i - \mathbf{a}_i^\top \mathbf{x})$:

- It removes the constraint, but modifies the cost.
- When $b_i - \mathbf{a}_i^\top \mathbf{x}$ is a very small positive number, $\log(b_i - \mathbf{a}_i^\top \mathbf{x})$ is a very big negative number, hence the minus sign in front.
- Barrier function does not behave well outside of the domain, when $b_i - \mathbf{a}_i^\top \mathbf{x} < 0$.

Hence the following QP:

$$\begin{aligned} \underset{\mathbf{x}}{\text{minimize}} \quad & \mathbf{x}^\top \mathbf{H}\mathbf{x} + \mathbf{f}^\top \mathbf{x}, \\ \text{subject to} \quad & \begin{cases} \mathbf{A}\mathbf{x} \leq \mathbf{b}, \\ \mathbf{C}(\mathbf{x}) = \mathbf{d}. \end{cases} \end{aligned} \tag{5}$$

...can be approximated as:

$$\begin{aligned} \underset{\mathbf{x}}{\text{minimize}} \quad & \mathbf{x}^\top \mathbf{H}\mathbf{x} + \mathbf{f}^\top \mathbf{x} - \sum_{i=1}^{n} \log(b_i - \mathbf{a}_i^\top \mathbf{x}), \\ \text{subject to} \quad & \mathbf{C}(\mathbf{x}) = \mathbf{d} \end{aligned} \tag{6}$$

Analytic center of linear inequalities

We can define *analytic center of linear inequalities* as a minimum of the function $J = -\sum\limits_{i=1}^{n} \log(b_i - \mathbf{a}_i^\top \mathbf{x})$. And that can be solved as a convex optimization:

$$\mathbf{x}_a = \underset{\mathbf{x}}{\operatorname{argmin}} \quad -\sum_{i=1}^{n} \log(b_i - \mathbf{a}_i^\top \mathbf{x})$$

At the analytic center of linear inequalities the shape of contour lines can be analysed as a local quadratic approximation of the function J:

$$\mathcal{C} = \{\mathbf{x} : \ (\mathbf{x} - \mathbf{x}_a)^\top \frac{\partial^2 J}{\partial \mathbf{x}^2} (\mathbf{x} - \mathbf{x}_a) = \epsilon\} \tag{7}$$

where ϵ is a small number.

Illustration of a barrier functions

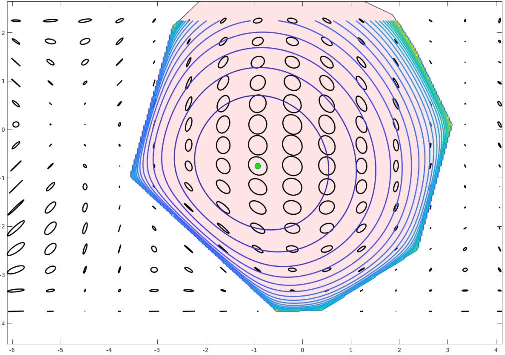

Figure 1: Barrier functions

Pink is the domain. The ellipsoids represent the shape of the hessian $\frac{\partial^2 J}{\partial \mathbf{x}^2}$ at different points on the domain. Green dot is $\mathbf{x}_a$.

Consider an optimization problem:

$$\begin{aligned} \underset{\mathbf{x}}{\text{minimize}} \quad & f_0(\mathbf{x}), \\ \text{subject to} \quad & \begin{cases} f_i(\mathbf{x}) \leq 0, \\ h_j(\mathbf{x}) = 0. \end{cases} \end{aligned} \tag{8}$$

It's *Lagrangian* is given as:

$$L(\mathbf{x}, \lambda_i, \nu_j) = f_0(\mathbf{x}) + \sum_i \lambda_i f_i(\mathbf{x}) + \sum_j \nu_j h_j(\mathbf{x}) \tag{9}$$

where λ_i and ν_j are Lagrange multipliers; they are sometimes called *dual variables*.

Lagrange dual function

Given *Lagrangian* $L(\mathbf{x}, \lambda_i, \nu_j) = f_0(\mathbf{x}) + \sum_i \lambda_i f_i(\mathbf{x}) + \sum_j \nu_j h_j(\mathbf{x})$, the associated *Lagrange dual function* is given as:

$$g(\lambda_i, \nu_j) = \inf_{\mathbf{x}} L(\mathbf{x}, \lambda_i, \nu_j). \tag{10}$$

Lagrange dual function is always concave. If p^* is the optimal value of the cost function of the original problem, then $g(\lambda_i, \nu_j)$ gives as a *lower bound* on its possible values. In fact, substituting any ν_j and $\lambda_i > 0$ gives us a valid lower bound on the cost. Maximum of $g(\lambda_i, \nu_j)$ over the domain given by $\lambda_i > 0$ provides us optimal (largest) lower bound of the problem, denoted as g^*.

Duality gap, strong and weak duality

If p^* is the optimal value of the cost function of the original problem and g^* is the optimal lower bound of the problem, then $p^* - g^*$ is called optimal *duality gap*.

If optimal duality gap is zero, the problem is said to have *strong duality*. If optimal duality gap greater than zero, the problem is said to have *weak duality*.

Lagrange dual function for a QP, 1

Consider the following QP:

$$\begin{aligned} \underset{\mathbf{x}}{\text{minimize}} \quad & \mathbf{x}^\top \mathbf{H} \mathbf{x}, \\ \text{subject to} \quad & \mathbf{A}\mathbf{x} \leq \mathbf{b}. \end{aligned} \tag{11}$$

Its Lagrangian is:

$$L(\mathbf{x}, \lambda) = \mathbf{x}^\top \mathbf{H} \mathbf{x} + \lambda^\top (\mathbf{A}\mathbf{x} - \mathbf{b}) \tag{12}$$

In order to minimize the Lagrangian with respect to $\mathbf{x}$ we find the gradient and set it to zero:

$$\frac{\partial L(\mathbf{x}, \lambda)}{\partial \mathbf{x}} = 2\mathbf{x}^\top \mathbf{H} + \lambda^\top \mathbf{A} = 0 \tag{13}$$

With that we can compute $\mathbf{x}$ as a function of λ:

$$\mathbf{x} = -0.5\mathbf{H}^{-1}\mathbf{A}^\top \lambda \tag{14}$$

Lagrange dual function for a QP, 2

Knowing that $\mathbf{x} = -0.5\mathbf{H}^{-1}\mathbf{A}^\top\lambda$ we can compute $g(\lambda)$ by substituting the $\mathbf{x}$ we found into the Lagrangian:

$$g(\lambda) = \frac{1}{4}\lambda^\top\mathbf{A}\mathbf{H}^{-1}\mathbf{H}\mathbf{H}^{-1}\mathbf{A}^\top\lambda - \frac{1}{2}\lambda^\top\mathbf{A}\mathbf{H}^{-1}\mathbf{A}^\top\lambda - \lambda^\top\mathbf{b} \quad (15)$$

$$g(\lambda) = -\frac{1}{4}\lambda^\top\mathbf{A}\mathbf{H}^{-1}\mathbf{A}^\top\lambda - \lambda^\top\mathbf{b} \quad (16)$$

In order to find the optimal lower bound we solve the following problem:

$$\begin{aligned} \underset{\lambda}{\text{maximize}} \quad & -\frac{1}{4}\lambda^\top\mathbf{A}\mathbf{H}^{-1}\mathbf{A}^\top\lambda - \lambda^\top\mathbf{b}, \\ \text{subject to} \quad & \lambda \geq 0. \end{aligned} \quad (17)$$

Note that optimal values of λ determine local sensitivity of the system with respect to small perturbations of constraints.

Consider minimizing $(\mathbf{x}-\mathbf{c})^\top(\mathbf{x}-\mathbf{c})$ when the domain is the second quadrant: $x_1 \geq 0$ and $x_2 \leq 0$. Find sensitivity of the problem as a function of $\mathbf{c}$.

$$\begin{aligned} \underset{\mathbf{x}}{\text{minimize}} \quad & (\mathbf{x}-\mathbf{c})^\top(\mathbf{x}-\mathbf{c}), \\ \text{subject to} \quad & \mathbf{A}\mathbf{x} \leq 0. \end{aligned} \tag{18}$$

where $\mathbf{A} = \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix}$.

The dual Lagrange function is:

$$g(\lambda) = -\frac{1}{4}\lambda^\top \mathbf{A}\mathbf{A}^\top \lambda + \lambda^\top \mathbf{A}\mathbf{c} \tag{19}$$

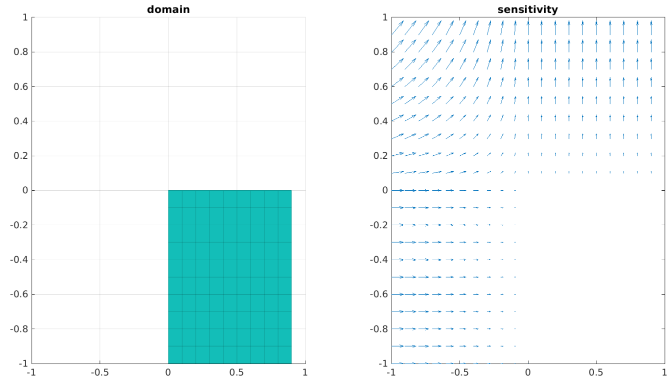

Figure 2: Sensitivity

Turquoise on the left is the domain. The arrows on the right show the values of λ.

Visualize contours of a quadratic program of your choice.
Compute its optimal lower bound and duality gap.

Lecture slides are available via Moodle.

You can help improve these slides at:
github.com/SergeiSa/Computational-Intelligence-Slides-Spring-2022

Check Moodle for additional links, videos, textbook suggestions.