UNIVERSITY OF CALIFORNIA, BERKELEY

Lecture 15 - EE127 Complementary Slackness and KKT

Oscar Ortega

July 16, 2021

1 REVIEW AND REFINEMENTS

Recall the definition given of Slater's Sufficient condition for Strong Duality: Given a the following optimization problem:

$$\min f_0(x)$$

$$f_i(x) \leq 0, i = 1, 2, ..., m$$

$$Ax = b$$

Then Slater's condition states that if there exists $x_0 \in (D)$ s.t $f_i(x_0) < 0, i = 1, ..., m, h_i(x_0) = 0, i = 1, ..., n$ Then strong duality will hold.
We can relax the following condition to the following and have strong duality still hold. Here is the relaxed sufficient condition:

$$\min f_0(x) \in \text{relint}(D)$$

$$\text{s.t } f_i(x^*) < 0 i = 1, 2, ..., m$$

$$Ax^* = b$$

In other words, x no longer has to be in strict interior of the domain but only needs to be in the relative interior of the domain, and we can now activate Linear Constraints. Below is the definition of the relative interior of a convex set S:

$$\text{relint}(S) := \{x \in S : \forall y \in S \exists \lambda > 1 : \lambda x + (1 - \lambda) y \in S\}$$

2 LASSO REGULARIZATION

Consider the following minimization problem:

$$\min_x \frac{1}{2}\|Ax-b\|_2^2 + \mu\|x\|_1, \mu \geq 0$$

Note how we can't explicitly take the dual since there are no explicit constraints? How would we fix this? We can reparametrize. This means we can now redefine the minimization as follows:

$$\min_{x,z}\|z\|_2^2 + \mu\|x\|_1$$

$$\text{s.t } Ax - b = z$$

Our Lagrangian is now the following:

$$\mathscr{L}(z,x,v) = \frac{1}{2}\|z\|_2^2 + \mu\|x\|_1 + v^T(z - Ax + b)$$

$$g(v) = \min_{x,z}\mathscr{L}(z,x,v)$$

$$= v^T b + \min_x(\mu\|x\|_1 - v^T Ax) + \min_z(\frac{1}{2}\|z\|^2 + v^T z)$$

$$= v^T b + \min_x(\mu\|x\|_1 - v^T Ax) + -\frac{1}{2}\|v\|^2$$

How do we minimize $\min_x(\mu\|x\|_1 - v^T Ax)$? Well, let's see what happens in the scalar case. this is equal to the following.

$$= \min_x(\mu|x| - v\alpha x)$$

$$g(x) = \mu|x| - v\alpha x$$

$$g(x) = \begin{cases} \mu x - v\alpha x & x \geq 0 \\ -\mu x - v\alpha x & x < 0 \end{cases}$$

$$g(x) = \begin{cases} (\mu - v\alpha)x & x \geq 0 \\ -(\mu + v\alpha)x & x < 0 \end{cases}$$

$$\min_x q(x) = \begin{cases} -\infty & \mu - v\alpha \leq 0 x \geq 0 \\ 0 & \mu - v\alpha \geq 0, x \geq 0 \\ -\infty & \mu + v\alpha > 0, x < 0 \end{cases}$$

$$\min_x q(x) = \begin{cases} -\infty & \mu \leq |v\alpha| \\ 0 & \mu > |v\alpha| \end{cases}$$

Which means our dual formulation becomes the following:

$$\max_v vb - \frac{1}{2}v^2$$

$$\text{s.t } |\nu\alpha| \leq \mu$$

If we now let $x \in \mathbb{R}^2$ our minimization becomes the following:

$$\min_{x_1}(\mu|x_1| - \nu_1\alpha_1 x_1) + \min_{x_2}(\mu|x_2| - \nu\alpha_1 x_2)$$

Our dual becomes the following:

$$d^* = \max_v v^T b - \frac{1}{2}\|v\|_2^2$$

$$\text{s.t } \alpha_1 v \leq \mu \text{ and } \alpha_2 v \leq \mu$$

Note that this is just a constraint to the l_∞ norm. So we can rewrite the dual as follows:

$$d^* = \max_v v^T b - \frac{1}{2}\|v\|_2^2$$

$$\|A^T v\|_\infty \leq \mu$$

This is why the l_∞ norm as a **dual norm** to the l_1 norm. ‘

3 Complementary Slackness and KKT conditions

$$\text{Let } p^*, d^*, x^*, \nu^*, \lambda^* \text{ optimal}$$

$$\text{Let } f_0(x^*) = g(\lambda^*, \nu^*)$$

$$= \inf_x f_0(x) + \sum_{i=1}^{n} \lambda_i^* f_i(x) + \sum_{i=1}^{n} \nu_i^* h_i(x)$$

$$= f_0(x^*) + \alpha + 0 : \alpha \leq 0$$

$$\leq f_0(x^*)$$

Notice how we proved the following:

$$f_0(x^*) \leq \mathscr{L}(x^*, \lambda^*, \nu^*) \leq f_0(x^*)$$

$$\rightarrow f_0(x^*) = \mathscr{L}(x^*, \lambda^*, \nu^*) \text{ and } \sum_{i=1}^{n} \lambda_i^* f_i(x) = 0$$

This implies what is known as **complementary slackness** which tells us the following.

$$\lambda_i^* > 0 \rightarrow f_i(x^*) = 0, i = 1, \ldots n$$

$$f_i(x^*) < 0 \rightarrow \lambda_i^* = 0, i = 1, \ldots, n$$

4 KKT CONDITIONS

Assume that the functions $f_0, ..., f_m$ $h_1, ..., h_p$ are differentiable and therefore have open domains.

$$\text{Let } p^*, d^*, x^*, \nu^*, \lambda^* \text{ optimal}$$

Then it must be the case that the gradient of the Lagrangian must be 0.
This implies the following:

$$f_i(x^*) < 0, i = 1, ..., m$$

$$h_i(x^*) = 0, i = 1, ..., p$$

This is known as primal feasibility.

$$\lambda_i^* \geq 0, i = 1, ..., m$$

This is known as dual feasibility.

$$\lambda_i^* \geq 0, i = 1, ..., m$$

Recall, this condition is known as complementary slackness. As well as the condition described above, which we know as **stationarity** these are known as the **Karush-Kuhn-Tucker** (KKT) conditions. Furthermore, if if we have a convex primal problem, the KKT conditions are sufficient to assume strong duality.