– The condition can be sharpened:

Ex. the affine inequalities do not need to hold with strict inequality.

Relaxed Slater's constraint qualification

The strong duality holds for the convex problem if there exists $x^* \in \text{relint}(D)$ such that

$$f_i(x) \leq 0, \quad i = 1, \ldots, k$$

if $f_i(x)$, for $i = 1, \ldots, k$, are affine.

and $f_i(x) < 0$, $i = k+1, \ldots, m$,

$Ax = b$.

– The strong duality holds for any LP provided at least one of the primal or dual problems is feasible.

Optimality conditions

- A dual feasible (λ, ν) establishes a lower bound on the optimal value of the primal problem: $p^* \geq g(\lambda, \nu)$.

- x is ϵ-suboptimal if $f_0(x) \leq p^* + \epsilon$
  (λ, ν) $g(\lambda, \nu) \geq d^* - \epsilon$.

- If x is primal feasible, (λ, ν) is dual feasible then $\epsilon \triangleq f_0(x) - g(\lambda, \nu)$ is the duality gap associated with the points x and (λ, ν).

$$p^* \in [g(\lambda, \nu), f_0(x)], \quad d^* \in [g(\lambda, \nu), f_0(x)].$$

x, and (λ, ν) are ϵ-suboptimal.

- If the duality gap is zero, i.e.
  $f_0(x) = g(\lambda, \nu) \Rightarrow$ x is primal optimal. (λ, ν) is dual optimal.

(λ, ν) is a certificate that x is primal optimal and vice versa.

Example: Quadratic program

$$\text{minimize } x^T x \quad \text{s.t. } Ax = b. \iff \text{maximize } -\frac{1}{4}\nu^T A A^T \nu - b^T \nu.$$

Slater's condition is that the primal problem is feasible. (No. inequality constraint).

Thus $p^* = d^*$ if $b \in R(A)$.

In fact, strong duality holds for any optimization problem with quadratic objective and one quadratic inequality constraint provided Slater's condition holds.

(See Appendix B if you are interested).

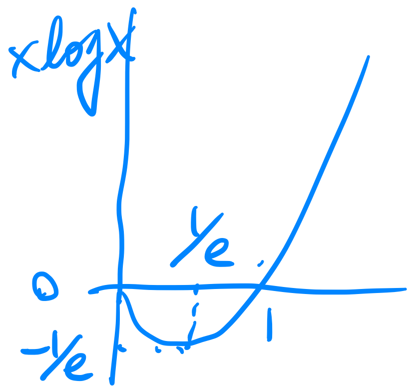
Ex. Entropy maximization Problem

$$\text{minimize} \quad \sum_{i=1}^{n} x_i \log x_i$$
$$\text{s.t.} \quad Ax \le b$$
$$\sum_{i=1}^{n} x_i = 1,$$
$$x > 0$$

a convex problem

(In Information Theory.
$H(\{p_i\}) = -\sum_{i=1}^{n} p_i \log p_i$
Shannon entropy.)

$f_0(x) = \sum_{i=1}^{n} x_i \log x_i$ is convex



$$\frac{\partial f_0(x)}{\partial x_i} = \log x_i + 1.$$

$$\frac{\partial^2 f_0(x)}{\partial x_i x_j} = 0. \quad \frac{\partial^2 f_0(x)}{\partial x_i^2} = \frac{1}{x_i} > 0 \quad \text{for } x_i > 0.$$

The Lagrangian $L(x, \lambda, \nu) = \sum_{i=1}^{n} x_i \log x_i + \lambda^T(Ax - b) + \nu(x^T \mathbf{1} - 1).$

$\nu \in \mathbb{R}$. $\lambda \in \mathbb{R}^m$

$$0 = \frac{\partial L}{\partial x_i} = \log x_i + 1 + \sum_{j=1}^{n} \lambda_j A_{ji} + \nu$$

$$\Rightarrow \quad x_i = e^{-\nu - 1 - \sum_{j=1}^{n} \lambda_j A_{ji}}$$

$$\text{maximize} \quad -b^T\lambda - \underline{\underline{\nu}} - e^{-\nu-1} \sum_{i=1}^{n} e^{-\sum_{j=1}^{n} \lambda_j A_{ji}} = \tilde{f}_0(\lambda, \nu)$$
$$\text{s.t.} \quad \underline{\underline{\lambda \ge 0}}$$

$$\Downarrow$$

$$0 = \frac{\partial \tilde{f}_0(\lambda, \nu)}{\partial \nu} \Rightarrow \nu = \log \sum_{i=1}^{n} e^{-\sum_{j=1}^{n} \lambda_j A_{ji}} - 1.$$

Suppose an algorithm produces a sequence of primal feasible $x^{(k)}$ and dual feasible $(\lambda^{(k)}, \nu^{(k)})$ for $k=1,2,\dots$

and $\epsilon_{abs} > 0$ is a given required absolute accuracy. Then the stopping criterion $f_0(x^{(k)}) - g(\lambda^{(k)}, \nu^{(k)}) < \epsilon_{abs}$

guarantees that when the algorithm terminates, $x^{(k)}$ is ϵ_{abs}-suboptimal and $(\lambda^{(k)}, \nu^{(k)})$ is a certificate.

provided the strong duality holds.

Complementarity Slackness.

Assume that strong duality holds.

x^* is primal optimal, (λ^*, ν^*) is dual optimal.

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

by def.

$$= \inf_x \left(f_0(x) + \sum_{i=1}^{m} \lambda_i^* f_i(x) + \sum_{j=1}^{q} \nu_j^* h_j(x) \right)$$

$$\leq f_0(x^*) + \sum_{i=1}^{m} \underbrace{\lambda_i^*}_{\geq 0} \underbrace{f_i(x^*)}_{\leq 0} + \underbrace{\sum_{j=1}^{q} \nu_j^* h_j(x^*)}_{= 0}$$

$$\leq \boxed{f_0(x^*).}$$

The two inequalities hold with equality.

— x^* minimizes $L(x, \lambda^*, \nu^*)$

— $\sum_{i=1}^{m} \underbrace{\lambda_i^*}_{\geq 0} \underbrace{f_i(x^*)}_{\leq 0} = 0.$

$\Leftrightarrow \quad \lambda_i^* f_i(x^*) = 0 \quad$ for $i = 1, \dots, m$

$$\begin{cases} \lambda_i^* > 0 \Rightarrow f_i(x^*) = 0. \\ f_i(x^*) < 0 \Rightarrow \lambda_i^* = 0. \end{cases}$$

(complementary slackness).

$$\text{maximize} \quad -b^T\lambda - \log \sum_{i=1}^{n} e^{-\sum_{j=1}^{n} \lambda_j A_{ji}}$$

$$\text{s.t.} \quad \lambda \geq 0.$$

$\Rightarrow$ a geometric program. (GP).

Karush-Kuhn-Tucker (KKT) conditions

~~(Necessary~~ ~~for strong duality~~)

Assume that $f_0, f_1, \cdots, f_m, h_1, \cdots, h_q$ are differentiable.

Let x^* and (λ^*, ν^*) be any primal and dual optimal points with zero duality gap,

1. $f_i(x^*) \le 0, \ i=1, \dots, m$. ($x^*$ is feasible).
   $h_j(x^*) = 0, \ j=1, \dots q$.

2. $\lambda^* \ge 0$ (λ^* is feasible).

3. $\lambda_i^* f_i(x^*) = 0, \ i=1, \dots, m$.
   (complementary slackness).

4. $\nabla f_0(x^*) + \sum_{i=1}^{m} \lambda_i^* \nabla f_i(x^*) + \sum_{j=1}^{q} \nu_j^* \nabla h_j(x^*) = 0$.

Since x^* minimizes $L(x, \lambda^*, \nu^*)$ over x.

If the strong duality holds, and x, (λ, ν) are optimal, then they must satisfy these KKT conditions.

KKT conditions for $\boxed{\text{convex problem}}$

(also sufficient)

If $\tilde{x}, \tilde{\lambda}, \tilde{\nu}$ satisfy the KKT conditions for a convex problem, (f_i convex, h_j affine), then they are optimal with zero duality gap.

proof. ① $\tilde{x}$ is primal feasible.

② $L(x, \tilde{\lambda}, \tilde{\nu})$ is convex in x.

③ $\boxed{\nabla_x L(x, \tilde{\lambda}, \tilde{\nu}) = 0}$ at $x = \tilde{x}$.

$f_0(x)$ convex (unconstrained)

x^* is optimal $\Leftrightarrow \nabla f_0(x^*) = 0$

$$g(\tilde{\lambda}, \tilde{\nu}) = \inf_x \boxed{L(x, \tilde{\lambda}, \tilde{\nu})}$$

$$= f_0(\tilde{x}) + \underbrace{\sum_{i=1}^{m} \tilde{\lambda}_i f_i(\tilde{x})}_{=0} + \underbrace{\sum_{j=1}^{p} \tilde{\nu}_j h_j(\tilde{x})}_{=0}$$

$$= f_0(\tilde{x})$$

complementary slackness

$\tilde{x}$ is feasible.

necessary and sufficient conditions for convex problems

Ex. minimize $\frac{1}{2}x^T P x + q^T x + r.$

s.t. $Ax = b.$

where $P \in S_+^n$, $A \in \mathbb{R}^{m \times n}$.

$x \in \mathbb{R}^n$

$\nu \in \mathbb{R}^m$

The KKT conditions are

- $Ax^* = b$
- $Px^* + q + A^T \nu^* = 0.$

$$L(x, \nu) = \frac{1}{2}x^T P x + q^T x + r + \nu^T (Ax - b).$$

$$0 = \nabla_x L(x, \nu) = Px + q + A^T \nu$$

Solving these m+n equations in the m+n variables x^* and ν^*, gives the optimal primal and dual variables.

Solving the primal problem via the dual

Suppose the strong duality holds, and an optimal (λ^*, ν^*) is known.

Suppose the minimizer of $L(x, \lambda^*, \nu^*)$ is unique. If the solution to

$$\text{minimize} \quad f_0(x) + \sum_{i=1}^{m} \lambda_i^* f_i(x) + \sum_{j=1}^{q} \nu_j^* h_j(x)$$

is prime feasible. it must be primal optimal.