Duality 3 bis

$$\boxed{\begin{array}{l} \min \|\vec{x}-\vec{c}\|^2 \\ \text{subject to } Ax=b \end{array}}$$

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

A $m \times n$ real matrix fixed

$b \in \mathbb{R}^m$ fixed

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

(Example: $\vec{c} \in \mathbb{R}^2$ $\vec{c} = \begin{pmatrix} 7/2 \\ 2/7 \end{pmatrix}$

$$f(x) = \|x-c\|^2 = \left(x_1 - \frac{7}{2}\right)^2 + \left(x_2 - \frac{2}{7}\right)^2 \Big)$$

- $Ax=b$ represents m equality contraints on x:

$$(Ax)_j = b_j \quad , \quad j=1,2,\dots,m$$

Therefore, we introduce m Lagrange multipliers (or dual variables) $\mu_1, \mu_2, \dots, \mu_m$, and we construct the Lagrange function

$$\mathcal{L}(\vec{x}, \vec{\mu}) = f(x) - \sum_{j=1}^{m} \mu_j \left((Ax)_j - b_j\right) =$$

$$= \|x-c\|^2 - \langle \mu, Ax-b \rangle$$

$$= \langle x-c, x-c \rangle - \langle A^t\mu, x \rangle + \langle \mu, b \rangle$$

- We know that

$$\min_{\substack{\text{subject} \\ \text{to } Ax=b}} \langle x-c, x-c \rangle = \min_{x \in \mathbb{R}^n} \max_{\mu \in \mathbb{R}^m} \mathcal{L}(x,\mu)$$

- the duality gap is the difference:

$$DG = \min_{x \in \mathbb{R}^n} \max_{\mu \in \mathbb{R}^m} \mathcal{L}(x,\mu) - \max_{\mu \in \mathbb{R}^m} \min_{x \in \mathbb{R}^n} \mathcal{L}(x,\mu)$$

- In our case, there exist a saddle point and $DG = 0$

<u>because</u> → $\mathcal{L}(x,\mu)$ is convex on x, ($\forall \mu$ fixed), due to → (1)

→ $\mathcal{L}(x,\mu)$ is concave on μ, ($\forall x$ fixed), due to → (2)

① $\mathcal{L}(x,\mu) = \langle x-c, x-c\rangle - \langle A^t\mu, x\rangle + \langle \mu, b\rangle$

is convex on x because:

$$\nabla_x \mathcal{L}(x,\mu) = 2(x-c) - A^t\mu$$

$$D_x^2 \mathcal{L}(x,\mu) = 2I = \begin{pmatrix} 2 & & & 0 \\ & 2 & & \\ & & \ddots & \\ 0 & & & 2 \end{pmatrix}$$

positive definite matrix

$\Downarrow$

$\mathcal{L}$ convex on x (for every μ fixed)

② $\mathcal{L}(x,\mu) = \langle x-c, x-c\rangle - \langle \mu, Ax-b\rangle$

is linear on μ (for every x fixed)

Thus, $\mathcal{L}$ is ~~[illegible]~~ concave on μ.

- Thus, there exists a saddle point of $\mathcal{L}$ and $DG = 0$

thus, we can exchange min-max by max-min:

$$\min_{\substack{\text{subject to}\\ Ax=b}} \langle x-c, x-c\rangle = \min_x \max_\mu \mathcal{L}(x,\mu) = \max_\mu \min_x \mathcal{L}(x,\mu)$$

$$= \max_\mu \underbrace{\left[\min_x \mathcal{L}(x,\mu)\right]}_{g_D(\mu)} = \max_\mu g_D(\mu)$$

- Let's compute the dual function $g_D(\mu)$ by solving $\min_x \mathcal{L}(x,\mu)$ (for μ fixed). As $\mathcal{L}(x,\mu)$ is convex with respect to x and the minimization is over $\mathbb{R}^n$ $\Longrightarrow$ the necessary and sufficient condition of min is

$$\nabla_x \mathcal{L}(x,\mu) = 0$$

the necessary and sufficient condition is

$$\nabla_x \mathcal{L}(x,\mu) = 0$$
$$\parallel$$
$$2(x-c) - A^t\mu = 0$$
$$2x = A^t\mu + 2c \implies \boxed{x^*(\mu) = \tfrac{1}{2}A^t\mu + c}$$

- Then,

$$g_D(\mu) = \mathcal{L}(x^*(\mu),\mu) = \langle \tfrac{1}{2}A^t\mu + \cancel{c} - \cancel{c}, \tfrac{1}{2}A^t\mu + \cancel{c} - \cancel{c}\rangle - \langle A^t\mu, \tfrac{1}{2}A^t\mu + c\rangle + \langle \mu, b\rangle$$
$$= \tfrac{1}{4}\langle A^t\mu, A^t\mu\rangle - \tfrac{1}{2}\langle A^t\mu, A^t\mu\rangle - \langle A^t\mu, c\rangle + \langle \mu, b\rangle$$
$$= -\tfrac{1}{4}\langle A^t\mu, A^t\mu\rangle - \langle \mu, Ac\rangle + \langle \mu, b\rangle$$
$$= -\tfrac{1}{4}\|A^t\mu\|^2 - \langle \mu, Ac - b\rangle$$

- the Dual problem is $\max\limits_{\mu\in\mathbb{R}^m}\left(-\tfrac{1}{4}\|A^t\mu\|^2 - \langle \mu, Ac-b\rangle\right) =$

$$= \min_{\mu\in\mathbb{R}^m}\left(\tfrac{1}{4}\|A^t\mu\|^2 + \langle \mu, Ac-b\rangle\right) =$$

Observe that $\boxed{\max g_D = -\min(-g_D)}$

(and also that one, of course)

But this problem has solution because its objective function is convex.

Indeed:

$$= \min_{\mu\in\mathbb{R}^m}\underbrace{\left(\tfrac{1}{4}\langle AA^t\mu, \mu\rangle + \langle \mu, Ac-b\rangle\right)}_{-g_D(\mu)}$$

and $\nabla(-g_D)(\mu) = \tfrac{2}{4}AA^t\mu + Ac - b$

$\boxed{D^2(-g_D)(\mu) = \tfrac{1}{2}AA^t}$ which is a positive definite matrix

(indeed $\langle \tfrac{1}{2}AA^t x, x\rangle = \tfrac{1}{2}\langle A^t x, A^t x\rangle = \tfrac{1}{2}\|A^t x\|^2 \geq 0$ $\forall x$

Thus, the function is convex in $\mathbb{R}^m$ and the dual problem

$$\max_{\mu} g_D(\mu) = \min_{\mu}\left(-g_D(\mu)\right)$$

has a solution (at least) in $\mathbb{R}^m$. To compute it, we use the necessary and sufficient condition

$$\nabla(-g_D) = 0$$

$$\frac{1}{2} AA^t \mu + Ac - b = 0 \quad \Rightarrow \quad AA^t \mu = 2(b - Ac)$$

$$\boxed{\mu^* = 2(AA^t)^{-1}(b - Ac)}$$

optimal μ^*

To compute the primal solution of the primal problem $\min_{Ax=b} f(x)$ ~~...~~

we substitute $x^* = x^*(\mu^*) = \cancel{\frac{1}{2}} A^t \left[\cancel{2}(AA^t)^{-1}(b - Ac)\right] + c$

$$= A^t(AA^t)^{-1}(b - Ac) + c$$

Extra: let's verify that x^* satisfy the constraint $Ax = b$

$$Ax^* = A\left[A^t(AA^t)^{-1}(b - Ac) + c\right]$$

$$= \cancel{AA^t}\,\cancel{(AA^t)^{-1}}(b - Ac) + Ac$$

$$= b - \cancel{Ac} + \cancel{Ac} = b$$