一. 原问题与对偶问题

我的确理解错了原问题的概念，诚如您所说，下式才被称为原问题

$$\min_{\mathbf{w},b} \frac{1}{2}||\mathbf{w}||_2^2$$

$$s.t. \, y_i\left(\mathbf{w}\mathbf{x}_i + b\right) \geq 1 \qquad i = 1, \cdots, n$$

如果我们定义函数$\theta(\mathbf{w}, b) = \max\limits_{\alpha_i \geq 0} L(\mathbf{w}, b, \boldsymbol{\alpha})$，则$\theta$的极小化问题

$$\min_{\mathbf{w},b} \theta(\mathbf{w}, b) = \min_{\mathbf{w},b} \max_{\alpha_i \geq 0} L(\mathbf{w}, b, \boldsymbol{\alpha})$$

与原问题等价。

原问题的对偶问题为

$$\max_{\alpha_i \geq 0} \min_{\mathbf{w},b} L(\mathbf{w}, b, \boldsymbol{\alpha})$$

二. SVM推导过程中的KKT条件

在查看这篇文章后[支持向量机（SVM）——原理篇 - 知乎 (zhihu.com)]，我对其中的推导逻辑感到有些迷惑，这篇文章中说到

> 要有 $p^* = d^*$ ，需要满足两个条件：
>
> ① 优化问题是凸优化问题
>
> ② 满足KKT条件
>
> 首先，本优化问题显然是一个凸优化问题，所以条件一满足，而要满足条件二，即要求
>
> $$\begin{cases} \alpha_i \geq 0 \\ y_i\left(\boldsymbol{w_i} \cdot \boldsymbol{x_i} + b\right) - 1 \geq 0 \\ \alpha_i\left(y_i\left(\boldsymbol{w_i} \cdot \boldsymbol{x_i} + b\right) - 1\right) = 0 \end{cases}$$

并且后面也说到

> 前面的推导都是假设满足KKT条件下成立的，KKT条件如下
>
> $$\begin{cases} \alpha_i \geq 0 \\ y_i\left(\boldsymbol{w_i} \cdot \boldsymbol{x_i} + b\right) - 1 \geq 0 \\ \alpha_i\left(y_i\left(\boldsymbol{w_i} \cdot \boldsymbol{x_i} + b\right) - 1\right) = 0 \end{cases}$$

似乎KKT条件是我人为地让它满足，而非直接证明原问题满足KKT条件。因此我比较好奇，KKT条件是我对原问题附加的条件，还是原问题可以推导出其满足KKT条件？如果是后者，那么$\alpha_i\left[y_i\left(\mathbf{w}^T\mathbf{x}_i + b\right) - 1\right] = 0$也就是KKT条件中的$\alpha_i g_i(\mathbf{x}) = 0$，又该如何证明成立呢？