1. 证明：范数 $\|\cdot\|$ 的 1 对偶范数满足范数的定义

$\|z\|_* = \sup\{z^T\text{x}:\|x\| \le 1\} = \sup\{z^T\text{x}:\|x\| = 1\}$

1. 证明：1）非负性，正定性：$\forall z = 0, sup\{z^T x\} = 0$ so $\|0\|^* = 0$ if $z \neq 0, \exists x,\ thatz^T x \neq 0$ so $\sup\{z^T\text{x}:\|x\| \le 1\} \ge 0$

所以 $\forall x \in \mathbb{R}^n, \|x\|_* \ge 0$

同时，如果 $\|x\|_* = 0, x = 0$

2）齐次性：我们需要证明 $\|\alpha z\|_* = \|\alpha\|\|z\|_*$

$\|\alpha z\|_* = sup\{(\alpha z)^T x : \|x\| \le 1\} = sup\{\alpha z^T x : \|x\| = 1\}$

因为 α 是标量，所以可以将 α 提出到括号外，于是有

$\|\alpha z\|_* = sup\{\alpha z^T x : \|x\| = 1\} = \alpha sup\{z^T x : \|x\| = 1\} = \alpha\|z\|_*$

齐次性成立

3）三角不等式：我们需要证明 $\|z_1 + z_2\|_* \le \|z_1\|_*\|z_2\|_*$

$\|z_1 + z_2\|_* = sup\{(z_1 + z_2)^T x : \|x\| = 1\} = sup\{z_1^T x + z_2^T x : \|x\| = 1\}$ 根据上确界的性质，有 $sup\{z_1^T x + z_2^T x : \|x\| = 1\} \le sup\{z_1^T x : \|x\| = 1\} + sup\{z_2^T x : \|x\| = 1\} = \|z_1\|_*\|z_2\|_*$

三角不等式证毕

综上, 对偶范数满足范数的定义

2. 证明：$z^T\text{x} \le \|x\|\|z\|$

2. 证明: 因为 $\|z\|_* = sup\{z^T x : \|x\| \le 1\} = sup\{z^T x : \|x\| = 1\}$

所以 $\|z\|_* \ge z^T \frac{x}{\|x\|}$(当 z 与 x 同方向时，取"=")

对等式两边同时乘 $\|x\|$ 我们得到：$z^T x \le \|x\|\|z\|_*$ 证毕

3.$f : \mathbb{R}^n \to \mathbb{R}$, $f(x) = log \sum_{i=1}^m exp(a_i^T x + b_i)$, 其中 $a_1, a_2, \cdots\ a_m \in \mathbb{R}^n$, $b_1, b_2, \cdots\ b_m \in \mathbb{R}$

1). 求函数 f 的梯度 $\nabla f(x)$

2). 求函数 f 的二阶导数 $\nabla^2 f(x)$

3. 解：1) 设 $u = \sum_{i=1}^m exp(a_i^T x + b_i)$ 则 $f(x) = logu$, $\nabla f(x) = \frac{1}{u} * \nabla u$

$\nabla u = \sum_{i=1}^m \nabla exp(a_i^T x + b_i) = \sum_{i=1}^m exp(a_i^T x + b_i)a_i$

故 $\nabla f(x) = \frac{1}{\sum_{i=1}^m exp(a_i^T x+b_i)} \sum_{i=1}^m exp(a_i^T x + b_i)a_i$

2).$\nabla^2 f(x) = \frac{1}{u}\sum_{i=1}^m \exp(a_i^T x + b_i)a_i a_i^T - \frac{1}{u^2}\left(\sum_{i=1}^m \exp(a_i^T x + b_i)a_i\right)\left(\sum_{i=1}^m \exp(a_i^T x + b_i)a_i\right)^T$

where $u = \sum_{i=1}^m exp(a_i^T x + b_i)$

4. 令 $A \in \mathbb{R}^m * n, B \in \mathbb{R}^p * n, b \in \mathbb{R}^m, d \in \mathbb{R}^p$ 集合

$P = \{x \in \mathbb{R}^n : Ax \le b\ Bx = d\}$ 是凸集吗? 为什么?

4. 解：集合 $P = \{x \in \mathbb{R}^n : Ax \le b\ Bx = d\}$ 是凸集，理由如下：

对于 x_1, x_2 满足 $Ax_1 \le b$ 并且 $Ax_2 \le b, \forall\lambda \in [0, 1], A(\lambda x_1 + (1-\lambda)x_2) = \lambda Ax_1 + (1-\lambda)Ax_2 \le \lambda b + (1-\lambda)b = b$

故满足不等式约束的集合是凸集

对于 x_1, x_2 满足 $Bx_1 = d$ 并且 $Bx_2 = d, \forall\lambda \in [0, 1], B(\lambda x_1 + (1-\lambda)x_2) = \lambda Bx_1 + (1-\lambda)Bx_2 =$

$\lambda d + (1-\lambda)d = d$

故满足等式约束的集合是凸集综上 P 是凸集

5. 证明：最大值函数 $f(x) = max\{x_1, x_2, \cdots, x_n\}, x = [x_1, x_2, \cdots, x_n] \in \mathbb{R}^n$ 为凸函数

5. 证明：$\forall x, y \in \mathbb{R}^n, \forall \lambda \in [0,1], f(\lambda x + (1-\lambda)y) = max\{\lambda x + (1-\lambda)y\}$

根据最大值的性质，$max\{\lambda x+(1-\lambda)y\} \le max\{\lambda x\}+max\{(1-\lambda)y\} = f(\lambda x)+f((1-\lambda)y) = \lambda f(x) + (1-\lambda)f(y)$

即:$f(\lambda x + (1-\lambda)y) \le \lambda f(x) + (1-\lambda)f(y)$

故 $f(x)$ 是凸函数

6. 二次函数 $f(x) = \frac{1}{2}x^T Qx + p^T x + r, x \in \mathbb{R}^n$ 是否为凸函数？是否为凹函数？是否为严格凸函数？是否为严格凹函数?

解：$\nabla^2 f(x) = Q$, 由凸函数的二阶判定得：

如果 Q 为对称半正定矩阵，$f(x)$ 是凸函数;

如果 Q 为对称半负定矩阵，$f(x)$ 是凹函数;

如果 Q 为对称正定矩阵，$f(x)$ 是严格凸函数;

如果 Q 为对称负定矩阵，$f(x)$ 是严格凹函数;

7. 对于凸优化问题

$min\, f(x)$

$s.t.\ x \in X$

如果目标函数 f 是可微的，那么可行解 x^* 是最优解当且仅当 $\forall y \in X$ 有 $\nabla f(x^*)^T(y-x^*) \ge 0$

解：$\Leftarrow$：当 $\forall y \in X, \nabla^T f(x^*)(y-x^*) \ge 0$ 为最优解时，因为 $f(x)$ 是凸函数，凸函数有 $\forall x, y \in X, f(y) \ge f(x) + \nabla^T f(x)(y-x)$

所以 $\forall y \in X, f(y) \ge f(x^*) + \nabla^T f(x^*)(y-x^*)$

即 $\forall y \in X, f(y) - f(x^*) \ge \nabla^T f(x^*)(y-x^*) \ge 0$

即 $\forall y \in X, f(y) \ge f(x^*)$, $x*$ 为最优解

$\Rightarrow$：若 x^* 是最优解，但不满足 $\forall y \in X, \nabla f(x^*)^T(y-x^*) \ge 0$ 即 $\exists y_0 \in X, \nabla f(x^*)^T(y-x^*) < 0$

设 $z = \lambda x^* + (1-\lambda)y_0, \lambda \in [0,1]$, 当 $\lambda = 0, z = y_0, \nabla f(x^*)^T(z-x^*) < 0$, 那么 $\exists \lambda \in [0,1) that\, f(z) = f(\lambda x^* + (1-\lambda)y_0) < f(x^*)$，这与 x^* 是最优解矛盾，所有假设不成立

综上，可行解 x^* 是最优解当且仅当 $\forall y \in X$ 有 $\nabla f(x^*)^T(y-x^*) \ge 0$

8. 证明 $x^* = (1, \frac{1}{2}, -1)^T$ 是如下优化问题的最优解 $min \frac{1}{2}x^T Qx + p^T x + r$

$s.t. -1 \le x_i \le 1, i = 1,2,3$ 其中，

$$
\text{P=}\begin{bmatrix} 13 & 12 & -2 \\ 12 & 17 & 6 \\ -2 & 6 & 12 \end{bmatrix} \text{,q=}\begin{bmatrix} -22 \\ -14.5 \\ 13 \end{bmatrix} \text{,r} = 1
$$

8. 证明：设 $f(x) = \frac{1}{2}x^TQx + p^Tx + r$, 则 $\nabla^2 f(x) = P$

$$
\text{P=}\begin{bmatrix} 13 & 12 & -2 \\ 12 & 17 & 6 \\ -2 & 6 & 12 \end{bmatrix}
$$

为正定矩阵，故 $f(x)$ 为凸函数且可微，显然不等式限制函数也是凸函数，有 $\nabla f(x^*)^T = Px^* + q =$

$$
\begin{bmatrix} 13 & 12 & -2 \\ 12 & 17 & 6 \\ -2 & 6 & 12 \end{bmatrix} \begin{bmatrix} 1 \\ \frac{1}{2} \\ -1 \end{bmatrix} + \begin{bmatrix} -22 \\ -14.5 \\ 13 \end{bmatrix} = \begin{bmatrix} -1 \\ 0 \\ 2 \end{bmatrix}
$$

$y - x^* =$

$$
\begin{bmatrix} y_1 + 1 \\ y_2 \\ y_3 - 2 \end{bmatrix}
$$

故 $\nabla f(x^*)^T(y - x^*) = 3 - y_1 + 2y_3 \ge 3 - 1 + 2 * 1 = 0$

由题 7 可知，当 x^* 满足 $\forall y \in X, \nabla^T f(x^*)(y - x^*) \ge 0$ 时 x^* 是最优解，故 $x^* = (1, \frac{1}{2}, -1)^T$ 是优化问题的最优解