最优化练习题

要求：请写出详细的推导过程或证明过程。

1. 计算如下向量函数$F(\alpha,\beta)$关于α和β的雅可比矩阵:

$$F(\alpha,\beta)=\begin{pmatrix}\mathbf{diag}(e^{-\frac{\alpha}{\varepsilon}})\mathbf{K}e^{-\frac{\beta}{\varepsilon}}\\ \mathbf{diag}(e^{-\frac{\beta}{\varepsilon}})\mathbf{K}^\top e^{-\frac{\alpha}{\varepsilon}}\end{pmatrix}. \tag{1}$$

2. 考虑优化问题

$$\min_x \quad f(x)=x_1^2+2x_2^2+4x_1+4x_2.$$

(a) 证明如果使用精确线搜索的最速下降梯度法从初始点$x^1=(0,0)^\top$求解该问题，产生的点列为

$$x^{k+1}=\left(\frac{2}{3^k}-2,\left(-\frac{1}{3}\right)^k-1\right)^\top.$$

(b) 这个迭代点列收敛到哪一个点？收敛速度是什么？

3. 考虑函数$f(x)=x_1^2+x_2^2$，以及迭代点列$x^k=(1+\frac{1}{2^k})(\cos k,\sin k)^T$，请说明

(a) $\{f(x^{k+1})\}$ 是否收敛？若收敛，给出收敛速度。

(b) $\{x^{k+1}\}$ 是否收敛？若收敛，给出收敛速度。

4. 考虑对称矩阵的特征值问题

$$\min_{x\in\mathbb{R}^n} \quad x^TAx, \quad \text{s.t.} \quad \|x\|_2=1,$$

其中$A^T=A$。试分析其所有的局部极小点、鞍点以及全局极小点

5. 设可微函数$f(x)$的定义域$\mathbf{dom}f=\mathbb{R}^n$，如果梯度$\nabla f(x)$是$L$-利普希茨连续，即存在$L>0$有

$$\|\nabla f(x)-\nabla f(y)\|\le L\|x-y\|, \quad \forall\, x,y\in\mathbf{dom}f, \tag{2}$$

证明函数$f(x)$有二次上界：

$$f(y)\le f(x)+\nabla f(x)^{\mathrm{T}}(y-x)+\frac{L}{2}\|y-x\|^2, \quad \forall\, x,y\in\mathbf{dom}f. \tag{3}$$

6. 设函数$f(x)$是$\mathbf{dom}f=\mathbb{R}^n$上的凸可微函数。假设梯度$\nabla f(x)$是$L$-利普希茨连续，即存在$L>0$有

$$\|\nabla f(x)-\nabla f(y)\|\le L\|x-y\|, \quad \forall\, x,y\in\mathbf{dom}f, \tag{4}$$

证明：$\nabla f(x)$ 有余强制性, 即对任意的$x,y\in\mathbb{R}^n$, 有

$$(\nabla f(x)-\nabla f(y))^{\mathrm{T}}(x-y)\ge\frac{1}{L}\|\nabla f(x)-\nabla f(y)\|^2.$$

7. 设可微函数$f(x)$ 的定义域为$\mathbb{R}^n$ 且存在一个全局极小点x^*, 若$f(x)$ 为梯度L -利普希茨连续的, 则对任意的x 有

$$\frac{1}{2L}\|\nabla f(x)\|^2\le f(x)-f(x^*)\le\frac{L}{2}\|x-x^*\|_2^2.$$

8. 假设可微函数$f(x)$的梯度$\nabla f(x)$是L-利普希茨连续，且f是可微强凸函数，即存在常数$m > 0$使得对任意$x, y \in \mathbb{R}^n$和$\theta \in (0,1)$，有：

$$f(\theta x + (1-\theta)y) \leq \theta f(x) + (1-\theta)f(y) - \frac{m}{2}\theta(1-\theta)\|x-y\|^2.$$

证明:

$$(\nabla f(x) - \nabla f(y))^{\mathrm{T}}(x-y) \geq \frac{mL}{m+L}\|x-y\|^2 + \frac{1}{m+L}\|\nabla f(x) - \nabla f(y)\|^2.$$

9. 假设函数$f(x)$可微且梯度$\nabla f(x)$是L-利普希茨连续. 从初始步长$t = \hat{t} > 0$开始考虑线搜索回退法. 如果不等式

$$f(x - t\nabla f(x)) < f(x) - \frac{t}{2}\|\nabla f(x)\|_2^2 \tag{5}$$

不满足，则缩小步长$t = \beta t$，其中$\beta \in (0,1)$，然后继续验证(5)是否满足. 证明回退法生成的步长满足$t \geq \min\{\hat{t}, \beta/L\}$.

10. 假设函数$f(x)$可微且梯度$\nabla f(x)$是L-利普希茨连续. 考虑无约束优化问题：

$$\min_{x\in\mathbb{R}^n} \; f(x).$$

- 写出求解该问题的梯度法。
- 假设$f(x)$为非凸函数，明确给出梯度法步长的取法并写出收敛性结果.
- 假设$f(x)$为凸函数，明确给出梯度法步长的取法并写出的收敛性(复杂度分析)结果.
- 假设$f(x)$为强凸函数，明确给出梯度法步长的取法并写出的收敛性(复杂度分析)结果.

11. （考虑可微函数$f(x): \mathbb{R}^n \to \mathbb{R}$和点列$\{x^i\}_{i=1}^{\infty}$。令$k \geq 2$，$s^{k-1} = x^k - x^{k-1}$ 以及$y^{k-1} = \nabla f(x^k) - \nabla f(x^{k-1})$。

   (a) 请分别计算如下两个优化问题的全局最优解:

$$\min_{\alpha\in\mathbb{R}} \quad \|\alpha y^{k-1} - s^{k-1}\|_2^2, \tag{6}$$

$$\min_{\alpha\in\mathbb{R}} \quad \|y^{k-1} - \alpha^{-1}s^{k-1}\|_2^2. \tag{7}$$

   (b) 请给出问题(6)和(7)构造合理性的一些直观解释。并根据它们构造求解优化问题$\min_x \; f(x)$的梯度算法。

12. 给定实对称矩阵$A \in \mathcal{S}^n$, $b \in \mathbb{R}^n$, $\Delta > 0$。考虑优化问题

$$\min_{x\in\mathbb{R}^n} \quad x^TAx + 2b^Tx, \quad \text{s.t.} \quad \|x\|_2 \leq \Delta.$$

给出并证明点$z \in \mathbb{R}^n$ 是该问题最优解的充要条件。

13. 给定$A(x) = A_0 + x_1A_1 + \cdots + x_nA_n$ 并且系数矩阵A_i 对称，令

$$f(x) = \lambda_{\max}(A(x)) = \sup_{\|y\|_2=1} y^{\mathrm{T}}A(x)y.$$

计算函数$f(x)$的一个次梯度.

14. 给定a_i和b_i，写出如下问题的一阶最优性条件:

$$\min_x \quad f(x) = \max_{i=1,\cdots,m}(a_i^{\mathrm{T}}x + b_i).$$

15. 设$f(x)$是适当且闭的强凸函数，其强凸参数为$\mu > 0$，$f^*(y)$是$f(x)$的共轭函数，则$f^*(y)$是梯度$\frac{1}{\mu}$-利普希茨连续的可微函数.

16. 考虑BFGS公式：

$$B_{k+1} = B_k - \frac{B_k s_k s_k^\top B_k}{s_k^\top B_k s_k} + \frac{y_k y_k^\top}{y_k^\top s_k}.$$

(a) 证明:

$$\mathrm{trace}(B_{k+1}) = \mathrm{trace}(B_k) - \frac{\|B_k s_k\|_2^2}{s_k^\top B_k s_k} + \frac{\|y_k\|_2^2}{y_k^\top s_k}.$$

(b) 证明：

$$\det(B_{k+1}) = \det(B_k)\frac{y_k^\top s_k}{s_k^\top B_k s_k}.$$

17. 假设x^*是最优化问题$\min_{x\in\mathbb{R}^n}\ f(x)$的最优解。假设函数$f(x)$二阶可微，海瑟矩阵$\nabla^2 f(x)$在$x^*$的一个领域内利普希茨连续且$\nabla^2 f(x^*)$严格正定。考虑牛顿算法

$$x_{k+1} = x_k + p, \quad p = -(\nabla^2 f(x_k))^{-1}\nabla f(x_k).$$

如果初始点x_0充分靠近x^*，证明迭代序列$\{x_k\}$收敛到x^*且收敛速度为二阶。

18. 设函数$f(x) = \|x\|^\beta$，其中$\beta > 0$为给定的常数。考虑步长为1的牛顿法对$f(x)$ 进行极小化，初值$x^0 \neq 0$。试证明：

(a) 若$\beta > 1$且$\beta \neq 2$，则x^k收敛到0 的速度为线性的；

(b) 若$0 < \beta < 1$，则牛顿法发散。

19. 给定向量$g \in \mathbb{R}^n$, 对称矩阵$B \in \mathbb{R}^{n\times n}$ 和信赖域半径Δ，考虑信赖域子问题：

$$\min_{p\in\mathbb{R}^n} \quad m(p) = g^\top p + \frac{1}{2}p^\top Bp, \quad \text{s.t.} \quad \|p\|_2 \le \Delta.$$

(a) 写出下面问题的最优解p^s：

$$p^s = \arg\min_{p\in\mathbb{R}^n} \quad g^\top p, \quad \text{s.t.} \quad \|p\|_2 \le \Delta.$$

(b) 写出下面问题的最优解τ：

$$\tau = \arg\min_{\tau\ge 0} \quad m(\tau p^s), \quad \text{s.t.} \quad \|\tau p^s\|_2 \le \Delta.$$

(c) 令科西步为$p^c = \tau p^s$。证明如下不等式：

$$m(0) - m(p^c) \ge \frac{1}{2}\|g\|_2 \min\left\{\Delta, \frac{\|g\|_2}{\|B\|_2}\right\}.$$

20. 给定连续可微函数$f(x)$ 和列向量$l, u \in \mathbb{R}^n$。考虑优化问题：

$$\min_{x\in\mathbb{R}^n} f(x), \ \text{s.t.}\ l \le x \le u. \tag{8}$$

(a) 给出点$x \in \mathbb{R}^n$到$[l, u]$的投影算子$P(x)$的显式表达式：

$$P(x) = \arg\min_y \|x - y\|_2^2, \ \text{s.t.}\ y \in [l, u].$$

(b) 对所有的$x \in \mathbb{R}^n$ 和任意$y \in [l, u]$，证明不等式$(P(x) - x)^\top(y - P(x)) \ge 0$成立。

(c) 对所有的$x, y \in \mathbb{R}^n$，证明不等式$(P(x) - P(y))^\top(x - y) \ge \|P(x) - P(y)\|_2^2$成立。

(d) 证明问题(8)的KKT条件等价于

$$x - P(x - \nabla f(x)) = 0.$$

(e) 写出求解问题(8)的投影梯度法，并给出一种全局收敛性结果。要求写清楚步长如何选取，以及保证该收敛性所需要的假设条件。

(f) 写出求解问题(8)的二次罚函数法。该二次罚函数是否可微？如果可微，给出其梯度的表达式。该二次罚函数是否二次可微？

(g) 写出求解问题(8)的增广拉格朗日函数法。要求写出: i) 具体的增广拉格朗日函数的表达式，并且该函数只有一个优化变量x。ii) 具体的乘子更新公式。

21. 给定对称矩阵$C \in \mathbb{R}^{n\times n}$，令$p$为正整数且$p \in [1, n]$，考虑优化问题:

$$\begin{aligned} \max_{Y\in\mathbb{R}^{n\times p}} \quad & f(Y) = \mathrm{Tr}(CYY^{\mathrm{T}}) \\ \text{s.t.} \quad & \mathrm{diag}(YY^{\mathrm{T}}) = \mathbf{1}, \end{aligned} \tag{9}$$

其中对于对称矩阵$X \in \mathbb{R}^{n\times n}$，$\mathrm{Tr}(X) = \sum_{i=1}^{n} X_{ii}, \mathrm{diag}(X) = [X_{11}, X_{22}, \cdots, X_{nn}]^{\mathrm{T}}$。

注：如果对矩阵形式相对不熟悉，可以先将变量写成向量形式: $Y = \begin{pmatrix} y_1^{\top} \\ \cdots \\ y_n^{\top} \end{pmatrix}$，其中$y_i \in \mathbb{R}^p$是列向量.

(a) 讨论可行点Y处线性无关约束品性(LICQ)是否满足？

(b) 写出目标函数$f(Y)$关于Y的梯度，估计梯度的利普希茨常数。

(c) 优化问题(9)的一阶最优性(KKT)条件。

(d) 二阶充分条件：假设在可行点Y处存在拉格朗日乘子满足KKT条件，如果拉格朗日函数的海瑟矩阵在临界锥上严格正定，则Y是问题(9)的严格局部极小点。请写出拉格朗日函数的海瑟矩阵(可以是矩阵形式或者算子形式)和临界锥具体形式。

(e) 假设Y是局部极小点，写出拉格朗日乘子的显式表达形式(只与C和Y有关)。

(f) 给定矩阵B，计算如下投影算子的最优解：

$$P(B) = \arg\min_{Y\in\mathbb{R}^{n\times p}} \|Y - B\|_F^2, \text{ s.t. } \mathrm{diag}(YY^{\mathrm{T}}) = \mathbf{1}. \tag{10}$$

(g) 请根据(10)写出求解问题(9)的投影梯度法，并参考无约束优化的线搜索条件给出一种步长选取条件。简要讨论该算法的全局收敛性以及保证该收敛性所需要的假设条件。

22. 给定矩阵$A \in \mathbb{R}^{m\times n}$和向量$b \in \mathbb{R}^m$，$\mu > 0$，考虑优化问题:

$$\min_x \quad \psi(x) = \mu\|x\|_1 + \frac{1}{2}\|Ax - b\|_2^2. \tag{11}$$

(a) 写出并验证$\psi(x)$的次梯度表达式。给出并证明问题(11)的最优性条件。

(b) 给出求解问题(11)的次梯度算法及其步长的取法，写出收敛性(复杂度分析)结果。

(c) 写出邻近算子的具体表达式:

$$\mathrm{prox}_t(y) = \arg\min_x \ t\|x\|_1 + \frac{1}{2}\|x - y\|_2^2.$$

(d) 给出求解问题(11)的近似点梯度算法及其步长的取法，写出收敛性(复杂度分析)结果。

(e) 给出求解问题(11)的Nesterov加速算法及其步长的取法，写出收敛性(复杂度分析)结果。

(f) 引入合适的拆分写出问题(11)的等价形式，写出求解该等价形式的增广拉格朗日函数法。要求明确写出如何求解增广拉格朗日函数子问题。

(g) 继上一小问: 写出求解该等价形式的交替方向乘子法(ADMM)。要求推导并明确写出每一个子问题的显式解。

(h) 写出问题(11)的对偶问题，写出求解该对偶问题的增广拉格朗日函数法。要求明确写出如何求解增广拉格朗日函数子问题。

(i) 继上一小问: 写出求解该对偶问题的交替方向乘子法(ADMM)。要求推导并明确写出每一个子问题的显式解。

23. 给定矩阵$A \in \mathbb{R}^{m\times n}$，向量$b \in \mathbb{R}^m$。证明以下两种情况中有且仅有一种成立：

(a) 存在某一个$x \ge 0$ 使得$Ax = b$.

(b) 存在某一个p使得$A^\top p \ge 0$且$b^\top p < 0$.

24. 令$x_{[i]}$ 为$x \in \mathbb{R}^n$ 的从大到小排列的第i 个分量，$0 < r \le n$是给定的正整数。定义函数

$$f(x) = x_{[1]} + x_{[2]} + \cdots + x_{[r]}.$$

(a) 证明$f(x)$ 是凸函数。

(b) 计算$f(x)$的次梯度。

25. 证明：d^*是信赖域子问题

$$\min_d \quad m(d) = f + g^\mathrm{T} d + \frac{1}{2} d^\mathrm{T} B d, \quad \text{s.t.} \quad \|d\|_2 \le \varDelta \tag{12}$$

的全局极小解当且仅当d^*是可行的且存在$\lambda \ge 0$使得

$$(B + \lambda I)d^* = -g, \tag{13a}$$

$$\lambda(\varDelta - \|d^*\|) = 0, \tag{13b}$$

$$B + \lambda I \succeq 0. \tag{13c}$$

26. 令$f(x)$为闭凸函数，

$$y \in \partial f(x) \Longleftrightarrow x \in \partial f^*(y) \Longleftrightarrow x^T y = f(x) + f^*(y).$$

27. 推导二阶锥$C = \{(x,t) \in \mathbf{R}^{n\times 1} | \|x\|_2 \le t\}$的投影算子.

28. 令$h(x)$为适当闭凸函数，定义邻近算子：

$$\operatorname{prox}_h(x) = \arg\min_u \left(h(u) + \frac{1}{2}\|u - x\|_2^2 \right).$$

(a) 证明Moreau分解

$$x = \operatorname{prox}_h(x) + \operatorname{prox}_{h^*}(x).$$

(b) 证明prox_h 是固定非扩张的

$$\|\operatorname{prox}_h(x) - \operatorname{prox}_h(y)\|_2 \le \|x - y\|_2.$$

(c) 令$f : \mathbb{R}^n \to \mathbb{R}$ 为梯度L_f-利普希茨连续函数。固定任意$t > L_f$，对于任何$x \in \operatorname{dom} h$ 和定义

$$x^+ = \operatorname{prox}_{h/t}\left(x - \frac{1}{t}\nabla f(x) \right),$$

我们有

$$f(x^+) + h(x^+) \le f(x) + h(x) - \frac{t - L_f}{2}\|x^+ - x\|_2^2.$$

29. 给定对称矩阵$C, A_i \in \mathbb{R}^{n\times n}$，向量$b \in \mathbb{R}^m$。考虑半定规划：

$$\begin{aligned} \min_X \quad & \langle C, X \rangle \\ \text{s.t.} \quad & \langle A_1, X \rangle = b_1, \\ & \ldots \\ & \langle A_m, X \rangle = b_m, \\ & X \succeq 0, \end{aligned} \tag{14}$$

其中$\langle C, X \rangle = \sum_{i,j} C_{ij} X_{ij}$，$X \succeq 0$指$X$为半正定矩阵。

(a) 写出问题(14)的对偶问题。

(b) 写出问题(14)的最优性(KKT)条件（需写出充要条件，注意分别写清楚假设条件）。

(c) 考虑直接求解问题(14):

  i. 写出增广拉格朗日函数算法，要求写出具体的增广拉格朗日函数形式和乘子更新格式；

  ii. 写出该增广拉格朗日函数(关于X)的梯度和海瑟矩阵，并计算梯度的利普希茨常数；

  iii. 写出极小化增广拉格朗日函数的投影梯度算法，给出步长选择并简要讨论该算法的收敛性（复杂度分析）结果。

(d) 考虑求解问题(14)的对偶问题:

  i. 写出交替方向乘子法(ADMM)，要求写出每一个子问题的最优解和乘子更新格式。

  ii. 写出增广拉格朗日函数算法，要求: 1) 将半定变量消掉，写出具体的增广拉格朗日函数形式和乘子更新格式。2) 计算该增广拉格朗日函数的梯度，并验证是否二次可微?

30. 假设$f(x)$是适当且闭的强凸函数, 其强凸参数为$\mu > 0$, h是闭凸函数，$A \in \mathbb{R}^{m\times n}$为实数矩阵. 考虑如下形式的问题：

$$\min_{x\in\mathbb{R}^n} \quad \psi(x) = f(x) + h(Ax). \tag{15}$$

(a) 令$f^*(y)$是$f(x)$的共轭函数，证明$f^*(y)$在全空间$\mathbb{R}^n$上有定义，且$f^*(y)$是梯度$\frac{1}{\mu}$-利普希茨连续的可微函数.

(b) 利用共轭函数写出(15)的拉格朗日对偶问题，要求该对偶问题由f和h共轭函数定义，且只有一个对偶变量z，记为：

$$\min_{z} \quad \phi(z) = f^*(-A^T z) + h^*(z) \tag{16}$$

(c) 写出求解(b)中对偶问题关于对偶变量z和函数$h^*(z)$做步长为t的近似点梯度算法，给出步长选择并简要讨论该算法的收敛性（复杂度分析）结果。

(d) 引入约束$y = Ax$，写出与问题(15)等价的约束优化问题：

$$\begin{aligned} \min \quad & f(x) + h(y), \\ \text{s.t.} \quad & y = Ax. \end{aligned} \tag{17}$$

进一步写出约束优化问题(17)的拉格朗日函数和增广拉格朗日函数

$$L(x, y, z) = f(x) + h(y) - z^{\mathrm{T}}(y - Ax),$$
$$L_t(x, y, z) = f(x) + h(y) - z^{\mathrm{T}}(y - Ax) + \frac{t}{2}\|y - Ax\|^2,$$

证明(c)中给出的迭代格式与如下算法等价

$$\begin{aligned} x^{k+1} &= \arg\min_x L(x, y^k, z^k), \\ y^{k+1} &= \arg\min_y L_t(x^{k+1}, y, z^k), \\ z^{k+1} &= z^k + t(Ax^{k+1} - y^{k+1}). \end{aligned} \tag{18}$$

(e) 假设$f(x)$只是闭凸函数，写出求解(b)中对偶问题关于对偶变量z和函数$\phi(z)$做步长为t的近似点算法，给出步长选择并简要讨论该算法的收敛性（复杂度分析）结果。

(f) 写出求解(17)的ALM算法，并证明该算法与(e)中给出的近似点算法等价。

31. 假设$f : \mathbb{R}^n \to \mathbb{R}$ 闭凸可微，且其梯度∇f 为L-Lipschitz 连续，即$\|\nabla f(x) - \nabla f(y)\| \le L\|x - y\|$；$h : \mathbb{R}^n \to (-\infty, +\infty]$ 为适当的闭凸函数。考虑复合优化问题

$$\min_{x} \quad \psi(x) = f(x) + h(x), \tag{19}$$

且假设$x^* = \arg\min_x \psi(x)$存在。

(a) 设$\psi(x)$是L_ψ-Lipschitz 连续。证明：存在常数G，对任意次梯度$g \in \partial\psi(x)$，有$\|g\| \le G$.

(b) 考虑次梯度法

$$x^{k+1} = x^k - t_k g^k, \qquad g^k \in \partial\psi(x^k), \quad t_k > 0.$$

假设存在常数$G > 0$ 使得$\|g^k\| \le G$，记$\hat{\psi}^k := \min_{0\le i\le k} \psi(x^i)$。证明对任意$k \ge 0$ 有

$$2\left(\sum_{i=0}^{k} t_i\right)\left(\hat{\psi}^k - \psi(x^*)\right) \ \le\ \|x^0 - x^*\|^2 + \sum_{i=0}^{k} t_i^2 G^2.$$

(c) 给出一种t_k的选择，证明$\hat{\psi}^k - \psi(x^*)$ 的收敛速度为$O(1/\sqrt{k})$。

(d) 对$v \in \mathbb{R}^n$，定义

$$\mathrm{prox}_{th}(v) := \arg\min_x \left\{ h(x) + \tfrac{1}{2t}\|x - v\|^2 \right\}.$$

证明：$\mathrm{prox}_h(v)$存在且唯一。

(e) 取常数步长$t \in (0, 1/L]$，定义$G_t(x) := \frac{1}{t}(x - \mathrm{prox}_{th}(x - t\nabla f(x)))$。证明：

$$G_t(x) = 0 \iff 0 \in \nabla f(x) + \partial h(x) \iff x \text{ 为}(19)\text{的最优解}.$$

(f) 定义迭代$x^{k+1} = \mathrm{prox}_{th}(x^k - t\nabla f(x^k))$，证明：对任意$y \in \mathbf{dom}\psi$ 与任意$k \ge 0$，有

$$\psi(x^{k+1}) \ \le\ \psi(y) \ +\ \frac{1}{2t}\left(\|y - x^k\|^2 - \|y - x^{k+1}\|^2\right).$$

(g) 证明

$$\psi(x^k) - \psi(x^*) \le \ \frac{1}{2tk}\|x^0 - x^*\|^2, \qquad k \ge 1.$$

(h) 证明

$$\min_{0\le i<k} \|G_t(x^i)\| \ \le\ \sqrt{\frac{2(\psi(x^0) - \psi^*)}{t\,k}}, \qquad k \ge 1.$$

32. 已知$W \in \mathbb{R}^{n\times n}$ 为严格正定矩阵，且满足割线方程$Ws = y$，由W 诱导的加权Frobenius 范数定义为

$$\|H\|_W = \|W^{1/2} H W^{1/2}\|_F.$$

给定对称矩阵$M \in \mathbb{R}^{n\times n}$，考虑如下约束优化问题：

$$\begin{aligned} \min_H \quad & \|H - M\|_W, \\ \text{s.t.} \quad & H = H^\top, \\ & Hy = s. \end{aligned}$$

请利用最优性条件推导该问题的最优解，并给出其显式表达式。

33. 设函数$f(x) = \inf_{u+v=x} (g(u) + h(v))$ 为g 与h 的卷积下确界，证明其共轭函数满足：

$$f^*(y) = g^*(y) + h^*(y).$$

34. 假设$h : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$为闭凸函数，证明对任意给定的$z \in \mathbb{R}^n$，邻近算子$\mathrm{prox}_h(z)$ 的解存在且唯一。

35. 请严格推导以下优化问题的显式解：

$$\begin{aligned} \min_x \quad & \frac{1}{2}\|x - z\|_2^2, \\ \text{s.t.} \quad & 1^\top x = 1, \\ & x \ge 0. \end{aligned}$$

如果解的表达式中涉及某参数（例如通过方程隐式定义），则必须完整推导该参数的具体形式。

36. 已知x^* 为等式约束优化问题$\min f(x)$, s.t. $c_i(x)=0, \forall i\in\mathcal{E}$ 的一个严格局部极小解，且满足KKT 条件，其对应的拉格朗日乘子为λ_i^*（$i\in\mathcal{E}$）。l_1 罚函数定义为

$$P(x,\sigma) := f(x) + \sigma\sum_{i\in\mathcal{E}}|c_i(x)|.$$

证明：当罚因子$\sigma > \bar{\sigma}$ 时，x^* 也为$P(x,\sigma)$ 的局部极小解。其中$\bar{\sigma} = \max\limits_{i\in\mathcal{E}}|\lambda_i^*|$.

37. 考虑等式约束优化问题$\min f(x)$, s.t. $c_i(x)=0, \forall i\in\mathcal{E}$ 的增广拉格朗日函数

$$L_\sigma(x,\lambda) = f(x) + \sum_{i\in\mathcal{E}}\lambda_i c_i(x) + \frac{\sigma}{2}\sum_{i\in\mathcal{E}}c_i^2(x),$$

其中$\mathcal{E}=\{1,2,\ldots,m\}$, f 和c_i 为$\mathbb{R}^n\to\mathbb{R}$ 函数。定义$\psi(\lambda) := \min\limits_x L_\sigma(x,\lambda)$.

(a) 证明：$\psi(\lambda)$ 为凹函数。

(b) 求出超微分$\partial\psi(\lambda) = \{g\in\mathbb{R}^m \mid \psi(\mu)\le\psi(\lambda) + g^\top(\mu-\lambda),\ \forall\mu\in\mathbb{R}^m\}$ 的一个元素。

38. 教材习题2.15。

39. 假设$A\in\mathbb{R}^{m\times n}$ 为行满秩矩阵且$m\ll n$。请推导求解Basis Pursuit主问题的如下ADMM 子问题最优解的具体形式:

$$\pi^{k+1} = \mathrm{argmin}_{\pi\in\mathbb{R}^n} L_\sigma(x^{k+1},\pi,y^k,s^k),$$

其中

$$L_\sigma(x,\pi,y,s) = \|x\|_1 + y^\top(A\pi-b) + s^\top(x-\pi) + \frac{\sigma}{2}\|A\pi-b\|_2^2 + \frac{\sigma}{2}\|x-\pi\|_2^2.$$

进一步请利用SMW公式将该最优解表达式中求逆复杂度降低到和$(AA^\top)^{-1}$的计算量同一个量级。

40. 教材习题2.15。

41. 请针对线性规划

$$\min_x c^\top x,\ \text{s.t. } Ax=b, x\ge 0,$$

写出PDHG和Chambolle-Pock算法的具体迭代格式，并在CPU和GPU上分别实现这两种算法，比较它们在不同规模问题上的运行时间和收敛性能。