2-无约束优化方法　选择题作业-复习题

题量：44 题 | 满分：100 分

一、单选题（共 44 题，100 分）

1. 在梯度下降法中，为什么通常选择负梯度作为搜索方向?

A 因为它是一个单位向量。

B 因为它是最陡的上升方向。

C 因为它是最陡的下降方向。

D 因为它总是指向最小值。

✓ 答案：C

解析：由于梯度方向是函数在某一点增长最快的方向，所以负梯度方向是函数在该点下降最快的方向，因此在梯度下降法中被选作搜索方向。

2. 在无约束优化中，若目标函数 $f(\boldsymbol{x})$ 在某点 $\boldsymbol{x}^{(k)}$ 处的梯度为 $\nabla f(\boldsymbol{x}^{(k)}) = [3, 2, -1]^T$，则以下哪种方向是下降方向?

A $[1, 1, 1]^T$

B $[-1, -1, -1]^T$

C $[0, 0, -1]^T$

D $[1, -1, 0]^T$

✓ 答案：B

解析：计算方向与梯度的点积：$[-1, -1, -1]^T \cdot [3, 2, -1]^T = -3 - 2 + 1 = -4 < 0$，因此是下降方向。

3. 在最速下降法中，每次选择的方向是哪个?

A 任意方向，只要不是梯度方向。

B 正梯度方向。

C 负梯度方向。

D 单位向量方向。

> **✓ 答案：C**
>
> **解析**：最速下降法的核心思想是选择负梯度方向作为搜索方向，因为这是在当前点下降最快的方向。

4. 已知目标函数 $f(\boldsymbol{x})$ 在点 $\boldsymbol{x}^{(k)}$ 处的梯度为 $\nabla f(\boldsymbol{x}^{(k)}) = [-2, 1]^T$。下面哪个方向不是下降方向?

A $[1, 1]^T$

B $[-1, -1]^T$

C $[1, 0]^T$

D $[1, -2]^T$

> **✓ 答案：B**
>
> **解析**：计算 $[-1, -1]^T \cdot [-2, 1]^T = 2 - 1 = 1 > 0$，因此不是下降方向。

5. 0.618 法中，初始搜索区间 $[a_0, b_0]$ 上的两个试探点 x_1 和 x_2 是如何确定的?

A $x_1 = a_0 + (b_0 - a_0) \times 0.382;\ x_2 = a_0 + (b_0 - a_0) \times 0.618$

B $x_1 = a_0 + (b_0 - a_0) \times 0.5;\ x_2 = a_0 + (b_0 - a_0) \times 0.5$

C $x_1 = a_0 + (b_0 - a_0) \times 0.25;\ x_2 = a_0 + (b_0 - a_0) \times 0.75$

D $x_1 = a_0 + (b_0 - a_0) \times 0.618;\ x_2 = a_0 + (b_0 - a_0) \times 0.618$

✓ 答案：A

解析：0.618 法利用黄金分割比例，在初始区间上分别取两个对称点 x_1 和 x_2，它们的位置分别为区间长度的 0.382 和 0.618 倍处。这是保证每次迭代保留区间长度为原长的 0.618 倍的关键设计。

6. 0.618 法中，若两个分点 $f(x_1) < f(x_2)$，则下一步的搜索区间会变成：

A $[a, x_2]$

B $[x_1, b]$

C $[x_2, b]$

D $[a, x_1]$

✓ 答案：A

解析：在 0.618 法中，若 $f(x_1) < f(x_2)$，则说明极值点更可能位于左半部分，因此新的搜索区间应为 $[a, x_2]$。反之若 $f(x_1) > f(x_2)$，则区间变为 $[x_1, b]$。

7. 0.618 法的迭代过程中，每次迭代保留的区间长度与原区间的比值是?

A 0.5

B 0.618

C 0.707

D 0.809

✓ **答案：B**

解析：0.618 法是一种黄金分割法，每次迭代保留的区间长度为原区间的 0.618 倍，这是算法设计的核心思想之一。

8. 设初始区间为 $[a, b]$，第一个试探点 x_1 的位置如何确定?

A $x_1 = a + 0.618 \cdot (b - a)$

B $x_1 = a + 0.5 \cdot (b - a)$

C $x_1 = b - 0.618 \cdot (b - a)$

D $x_1 = a - 0.382 \cdot (b - a)$

✓ **答案：C**

解析：0.618 法中，试探点按照黄金分割比例 0.618 来选取，第一个点 x_1 的位置为 $a + 0.382 \cdot (b - a)$。

9. 当 0.618 法的初始区间长度为 1，终止条件为区间长度小于某个给定精度 ε 时，迭代次数 n 应满足什么条件?

A $(0.618)^n < \varepsilon$

B $(0.618)^n > \varepsilon$

C $(0.5)^n < \varepsilon$

D $n < \log(\varepsilon)$

> **✓ 答案：A**
>
> **解析**：0.618 法的迭代过程每次将区间缩小至 0.618 倍，因此当区间长度 $L_n = L_0 \cdot (0.618)^n$ 小于 ε 时，迭代停止，即满足 $(0.618)^n < \varepsilon$。

10. 0.618 法的收敛速度属于哪种类型?

A 线性收敛

B 超线性收敛

C 二次收敛

D 三次收敛

> **✓ 答案：A**
>
> **解析**：0.618 法的收敛速度是线性的，因为每次迭代区间长度减少一个固定比例 0.618。

11. Wolfe 准则用于确定搜索方向的步长，其主要目的是保证什么?

A 目标函数数值单调递减

B 梯度方向与搜索方向一致

C 满足充分下降条件和曲率条件

D 步长不超过单位长度

✓ 答案：C

解析：Wolfe 准则通过两个条件（充分下降条件和曲率条件）确保在非精确线搜索中找到合适的步长，以平衡收敛速度和稳定性。

12. 在 Wolfe 准则中，下列哪个条件用于确保下降方向的有效性?

A $f(\boldsymbol{x}+\alpha \boldsymbol{d}) \leq f(\boldsymbol{x})+c_1 \alpha \nabla f(\boldsymbol{x})^T \boldsymbol{d}$

B $f(\boldsymbol{x}+\alpha \boldsymbol{d}) \geq f(\boldsymbol{x})+c_2 \alpha \nabla f(\boldsymbol{x})^T \boldsymbol{d}$

C $\nabla f(\boldsymbol{x}+\alpha \boldsymbol{d})^T \boldsymbol{d} \geq c_2 \nabla f(\boldsymbol{x})^T \boldsymbol{d}$

D $\alpha \leq 1$

✓ 答案：A

解析：这是充分下降条件，它确保函数值的减少足够大，从而保障搜索方向的有效性。

13. 下列关于 Wolfe 准则的描述，哪一项是正确的?

A 它是一种收敛性准则，用于保证算法的全局收敛性

B 它仅适用于二次优化问题

C 它是一种一维搜索准则，常用于线搜索方法中

D 它要求目标函数必须是凸函数

✓ 答案：C

解析：Wolfe 准则是一种在一维搜索中广泛使用的准则，用于确定步长选择的合理性。它确保下降方向上的步长既不会太小也不会太大。

14. 在优化问题中，Wolfe 准则主要用来：

A 求解最优解

B 决定搜索方向

C 确定步长

D 判断函数是否可微

✓ 答案：C

解析：Wolfe 准则是一维搜索中的常用准则，用于确定合适的步长。

15. 在最优化中，Armijo 准则与 Wolfe 准则之间有一个重要的区别。以下哪一项最准确地描述了这一区别?

A Armijo 准则不涉及曲率条件，而 Wolfe 准则涉及

B Wolfe 准则适用于凸函数，而 Armijo 准则适用于非凸函数

C Armijo 准则需要梯度信息

D Wolfe 准则总是比 Armijo 准则更高效

✓ 答案：A

解析：Armijo 准则只关注目标函数值的充分下降，而不考虑方向导数的变化率（即曲率条件）。而 Wolfe 准则除了包含 Armijo 准则的下降条件外，还引入了曲率条件，以进一步约束步长的选择。

16. 下列关于梯度法的说法，哪一个是正确的?

A 总是以最快的速度收敛

B 搜索方向是负梯度方向

C 适用于非光滑目标函数

D 收敛速度比牛顿法快

✓ 答案：B

解析：最速下降法的核心思想是沿着当前点的负梯度方向进行搜索。

17. 下列哪种方法最常用于确定最优化算法中的步长?

A 牛顿法

B 线搜索法

C 共轭梯度法

D 梯度法

✓ 答案：B

解析：线搜索法是最常用的步长选择方法，它在给定搜索方向的基础上寻找最优步长，使得目标函数值尽可能减小。

18. 梯度法中，目标函数 $f(\boldsymbol{x})$ 在第 k 步的更新公式是：

A $\boldsymbol{x}_{k+1} = \boldsymbol{x}_k + \alpha_k \nabla f(\boldsymbol{x}_k)$

B $\boldsymbol{x}_{k+1} = \boldsymbol{x}_k - \alpha_k \nabla f(\boldsymbol{x}_k)$

C $\boldsymbol{x}_{k+1} = \alpha_k \nabla f(\boldsymbol{x}_k)$

D $\boldsymbol{x}_{k+1} = \boldsymbol{x}_k \nabla f(\boldsymbol{x}_k)$

✓ 答案：B

解析：梯度法的迭代公式为 $x_{k+1} = x_k - \alpha_k \nabla f(x_k)$，其中 α_k 是步长，$\nabla f(x_k)$ 是梯度，搜索方向是负梯度方向。

19. 在梯度法中，搜索方向 d_k 是由以下哪一项决定的?

A 负梯度方向

B 梯度方向

C 正曲率方向

D 任意方向

✓ 答案：A

解析：在梯度法中，搜索方向 $d_k = -\nabla f(x_k)$，即沿着目标函数的负梯度方向进行移动，这是梯度下降法的核心思想。

20. 在使用 Python 实现梯度法时，常见的终止准则不包括以下哪项?

A 迭代次数达到上限

B 梯度范数小于阈值

C 函数值变化小于阈值

D 步长 α_k 小于某个固定值

✓ 答案：D

解析：步长 α_k 的大小受 Wolfe 条件影响，通常并不直接作为终止条件。常用终止条件是梯度范数、函数值变化或迭代次数限制。

21. 梯度法的收敛速度主要取决于目标函数的：

A 梯度大小

B Hessian 矩阵的条件数

C 初始点位置

D 函数的凸性

✓ 答案：B

解析：梯度法的收敛速度与目标函数的 Hessian 矩阵的条件数密切相关。如果 Hessian 矩阵的条件数较大，收敛速度会显著减慢。

22. 在使用牛顿法的过程中，若目标函数的 Hessian 矩阵非正定，则修正牛顿法采取的改进方法是____。

A 加入惩罚项

B 使用拟牛顿法

C 对 Hessian 矩阵进行修正以确保正定性

D 切换到梯度下降法

✓ 答案：C

解析：为了处理 Hessian 矩阵可能非正定的问题，牛顿法的变体通常引入正定修正项，以确保迭代方向始终有效。

23. 在修正牛顿法中，为了保证 Newton 方向的下降性质，通常通过什么方式修改 Hessian 矩阵?

A 添加一个小的单位矩阵

B 取 Hessian 的逆矩阵

C 用正定矩阵来替代原 Hessian 矩阵

D 将其与梯度向量相乘

✓ **答案：C**

解析：在修正牛顿法中，若 Hessian 矩阵不是正定的，可能导致搜索方向非下降。因此，常用的方法是用一个正定矩阵（如单位矩阵加上小扰动）代替原始的 Hessian 矩阵，以保证方向的下降性和算法的稳定性。

24. 在 Python 中实现修正牛顿法时，如果 Hessian 矩阵近似值不是正定的，以下哪种做法是合理的?

A 直接将其舍弃并采用梯度下降法

B 使用 Cholesky 分解进行处理

C 添加一个单位矩阵的倍数以保证正定性

D 忽略该问题继续运行

✓ **答案：C**

解析：在实际的 Python 代码中，当 Hessian 矩阵无法保证正定时，常用的修正方法是在其基础上加入一个小的正数乘以单位矩阵，从而构成一个正定矩阵。这种技术称为“正定化修正”，是修正牛顿法的核心思想之一。

25. 在修正牛顿法中，常用的正定修正形式是 $H_k + \mu I$，其中 $\mu > 0$ 的作用是什么?

A 增加计算效率

B 降低迭代次数

C 使矩阵正定，保证搜索方向的下降性

D 减小梯度模长

✓ 答案：C

解析：通过在 Hessian 矩阵中加入一个小的正数乘以单位矩阵，可以有效提升其正定性，从而确保修正后的 Hessian 矩阵能生成一个下降方向，这是修正牛顿法的关键策略。

26. 下列关于共轭方法的描述中，哪一项是正确的?

A 共轭方法是一种无需计算梯度的直接搜索方法。

B 共轭方法适用于任意非凸函数的最小化问题。

C 在正定二次目标函数中，共轭方法可以在 n 步内收敛到最优解。

D 共轭方法的搜索方向之间在欧几里得空间中正交。

✓ 答案：C

解析：共轭方向法适用于正定二次函数的最小化问题，在这种情况下，可以在 n 步内找到最优解。共轭方向并不是正交的，而是关于矩阵 H 共轭的。

27. 共轭方向法中，搜索方向满足什么条件?

A 搜索方向两两正交。

B 搜索方向两两共轭。

C 搜索方向两两平行。

D 搜索方向两两垂直。

✓ **答案：B**

解析：在共轭方向法中，搜索方向是关于矩阵 H 的共轭关系，即 $d_i^T H d_j = 0$，当 $i \neq j$ 时成立。

28. 设 A 是 $n \times n$ 对称正定矩阵，那么共轭向量集 $\{\boldsymbol{d}_1, \boldsymbol{d}_2, \ldots, \boldsymbol{d}_n\}$ 的数量最多是:

A $n-1$

B n

C $n+1$

D $2n$

✓ **答案：B**

解析：在 n 维空间中，最多可以构造 n 个关于 A 共轭的向量。这是由于共轭向量的定义依赖于线性无关性和共轭条件。

29. 以下哪种方法是共轭方向法的一个特例?

A 牛顿法

B 最速下降法

C 共轭梯度法

D 修正牛顿法

✓ 答案：C

解析：共轭梯度法（CG）是共轭方向法的一个重要特例，其搜索方向由梯度生成，并满足共轭条件。

30. 设 A 是一个对称正定矩阵，若 $d_1, d_2, \ldots, d_n$ 是一组关于 A 共轭的方向向量，则以下哪个结论成立?

A 这组向量线性相关。

B 这组向量构成 A 的特征向量集合。

C 这组向量是 A 的正交基。

D 这组向量是线性无关的。

✓ 答案：D

解析：共轭方向组中的方向向量是关于矩阵 A 共轭的，这意味着它们是线性无关的，但不一定正交。因此，选项 D 正确。

31. 共轭梯度法中的方向 d_k 需满足与之前的搜索方向 d_i $(i < k)$ 满足什么条件?

A 正交性

B 共轭性

C 相似性

D 线性相关性

✓ 答案：B

解析：共轭梯度法中的搜索方向需要满足共轭性 $d_k^T A d_i = 0,\ i < k$，其中 A 为正定二次目标函数的 Hessian 矩阵，这保证了算法在 n 步内收敛到最优解。

32. 在共轭梯度法中，搜索方向的更新公式通常依赖于当前的梯度和前一步的方向。下列哪个公式是 Fletcher-Reeves 版本中用于计算步长系数 β_k 的?

A $\beta_k = \dfrac{\|\boldsymbol{g}_k\|^2}{\|\boldsymbol{g}_{k-1}\|^2}$

B $\beta_k = \dfrac{\boldsymbol{g}_k^T \boldsymbol{g}_k}{\boldsymbol{g}_{k-1}^T A \boldsymbol{g}_{k-1}}$

C $\beta_k = \dfrac{\boldsymbol{g}_k^T (\boldsymbol{g}_k - \boldsymbol{g}_{k-1})}{\|\boldsymbol{g}_{k-1}\|^2}$

D $\beta_k = \dfrac{\boldsymbol{g}_k^T \boldsymbol{d}_k}{\boldsymbol{g}_{k-1}^T \boldsymbol{d}_{k-1}}$

✓ 答案：A

解析：Fletcher-Reeves 方法是共轭梯度法中的经典版本之一，其 β_k 系数的定义为当前梯度的模平方除以前一梯度的模平方。

33. 在使用共轭梯度法求解无约束优化问题 $\min f(\boldsymbol{x})$ 的 Python 实现中，下列哪个条件用于判断搜索方向 $\boldsymbol{d}_k$ 是否与 $\boldsymbol{d}_{k-1}$ 共轭?

A $\boldsymbol{d}_k^T A \boldsymbol{d}_{k-1} = 0$

B $\boldsymbol{d}_k^T \boldsymbol{g}_k = 0$

C $\boldsymbol{g}_k^T \boldsymbol{g}_{k-1} = 0$

D $\boldsymbol{x}_k^T \boldsymbol{x}_{k-1} = 0$

✓ 答案：A

解析：共轭梯度法的关键在于生成一组与矩阵 A 共轭的方向，即满足 $d_k^T A d_{k-1} = 0$。这是判断两个方向是否共轭的条件。

34. 在共轭梯度法中，如果目标函数为正定二次函数 $f(\boldsymbol{x}) = \frac{1}{2}\boldsymbol{x}^T A \boldsymbol{x} - \boldsymbol{b}^T \boldsymbol{x}$，则最多需要多少次迭代就可以达到最优解?

A n 次，其中 n 是变量维度

B n+1 次

C 无限次

D 取决于初始猜测

✓ 答案：A

解析：对于正定二次函数，共轭梯度法具有有限终止性质，最多迭代 n 次即可找到精确解，其中 n 是变量的维度。

35. 拟牛顿法中，为了保证迭代过程收敛，矩阵 H_k 应满足什么条件?

A 对称且正定

B 对称且半正定

C 正交

D 奇异

✓ 答案：A

解析：为了保证搜索方向是下降方向并确保算法收敛，矩阵 H_k 必须是对称正定的。

36. 拟牛顿法中，DFP 方法更新公式是基于哪种条件建立的?

A 梯度信息和函数值

B 搜索方向和步长

C 梯度差和位置差

D 海森矩阵的精确计算

✓ 答案：C

解析：DFP 方法基于梯度差 $y_k = g_{k+1} - g_k = \nabla f(x_{k+1}) - \nabla f(x_k)$ 和位置差 $s_k = x_{k+1} - x_k$ 构建更新公式。

37. 拟牛顿法中，BFGS 方法与 DFP 方法的主要区别在于?

A BFGS 方法直接修正 Hesse 矩阵，而 DFP 方法修正其逆矩阵

B BFGS 方法适用于非光滑函数，DFP 方法适用于光滑函数

C BFGS 方法无需计算逆矩阵

D DFP 方法收敛更快

> **✓ 答案：A**
>
> **解析**：BFGS 方法对 Hesse 矩阵进行直接修正，而 DFP 方法修正的是 Hesse 矩阵的逆矩阵。

38. 对称秩 1 算法（Symmetric Rank 1, SR1）在更新 Hesse 近似矩阵 B_{k+1} 时，其更新公式为：$B_{k+1} = B_k + \frac{(\boldsymbol{y}_k - B_k \boldsymbol{s}_k)(\boldsymbol{y}_k - B_k \boldsymbol{s}_k)^T}{(\boldsymbol{y}_k - B_k \boldsymbol{s}_k)^T \boldsymbol{s}_k}$。其中 $\boldsymbol{y}_k$ 和 $\boldsymbol{s}_k$ 分别代表什么?

A $\boldsymbol{y}_k = \boldsymbol{g}_{k+1} - \boldsymbol{g}_k$，$\boldsymbol{s}_k = \boldsymbol{x}_{k+1} - \boldsymbol{x}_k$

B $\boldsymbol{y}_k = \boldsymbol{s}_k$，$\boldsymbol{s}_k = \boldsymbol{g}_{k+1} - \boldsymbol{g}_k$

C $\boldsymbol{y}_k = H_k^{-1} \boldsymbol{g}_k$，$\boldsymbol{s}_k = \alpha_k \boldsymbol{d}_k$

D $\boldsymbol{y}_k = \boldsymbol{x}_k - \boldsymbol{x}_{k+1}$，$\boldsymbol{s}_k = \boldsymbol{g}_k$

> **✓ 答案：A**
>
> **解析**：对称秩 1 算法通过利用梯度的变化 $y_k = g_{k+1} - g_k$ 和迭代步长 $s_k = x_{k+1} - x_k$ 来更新 Hesse 近似矩阵。

39. 在 BFGS 算法中，Hesse 矩阵的更新方式是基于什么原则?

A 梯度方向的正交投影

B 拟牛顿条件和正定性保持

C 最小二乘法

D 共轭梯度法

✓ 答案：B

解析：BFGS 算法是一种拟牛顿方法，其核心思想是通过满足拟牛顿条件并维护 Hesse 矩阵的正定性来进行迭代优化。

40. BFGS 算法在每次迭代中通过哪个参数更新 Hesse 矩阵?

A 梯度差

B 搜索步长

C 梯度与变量的变化量

D 目标函数值的变化

✓ 答案：C

解析：BFGS 算法中，Hesse 矩阵的更新依赖于梯度差和变量差，这两个向量用于构造更新的校正项。

41. BFGS 算法的初始 Hesse 矩阵通常设为单位矩阵，这是为了保证什么特性?

A 计算的稳定性

B 正定性和对称性

C 收敛速度

D 搜索方向的确定

✓ 答案：B

解析：初始 Hesse 矩阵设为单位矩阵是为了满足正定性和对称性，便于后续的迭代更新。

42. DFP 算法中，Hesse 矩阵的更新公式依赖于哪两个变量?

A $\nabla f(\boldsymbol{x}_k)$ 和 $\nabla^2 f(\boldsymbol{x}_k)$

B $\boldsymbol{s}_k$ 和 $\boldsymbol{y}_k$

C $\boldsymbol{x}_k$ 和 $f(\boldsymbol{x}_k)$

D H_k 和 B_k

✓ 答案：B

解析：DFP 算法中，每次迭代都会根据搜索步长 $s_k = x_{k+1} - x_k$ 和梯度变化 $y_k = \nabla f(x_{k+1}) - \nabla f(x_k)$ 来更新 Hesse 的逆矩阵。

43. DFP 算法中，在初始化 Hesse 矩阵逆 H_0 时，通常选择哪种形式?

A 单位矩阵

B 零矩阵

C 随机矩阵

D 对角矩阵

✓ 答案：A

解析：DFP 算法通常以单位矩阵作为初始的 Hesse 逆矩阵 H_0。

44. DFP 算法中，如何确定搜索方向 d_k?

A $\boldsymbol{d}_k = -\nabla f(\boldsymbol{x}_k)$

B $\boldsymbol{d}_k = -H_k \nabla f(\boldsymbol{x}_k)$

C $\boldsymbol{d}_k = -B_k \nabla f(\boldsymbol{x}_k)$

D $\boldsymbol{d}_k = -\boldsymbol{s}_k$

✓ 答案：B

解析：DFP 算法中的搜索方向 d_k 是由负梯度乘以 Hesse 逆矩阵得到的，即 $d_k = -H_k \nabla f(x_k)$。