博士资格考试：数值计算（参考样卷）

试卷内容: 试卷包含两部分：第一部分是数学基础（70分），包含《偏微分方程数值方法》、《数值代数》和《数值分析》，是必答题；第二部分专业基础（30分），根据专业方向和招生导师的要求选做如下五个模块之一：《有限元方法》、《样条函数》、《小波分析》、《统计机器学习理论》和《计算机算法理论》。

答题要求: 所有题目的解答要有详细过程，用到的定理需要注明。

1. 数学基础（70分）

1.1 偏微分方程数值方法（30分）

一、考虑偏微分方程组

$$\mathbf{U}_t + \mathbf{A}\mathbf{U}_x = \mathbf{0}, \quad \mathbf{U} \in \mathbb{R}^m$$

其中，$\mathbf{A} \in \mathbb{R}^{m\times m}$是给定的实数矩阵，其特征值都是实数，且特征向量线性无关。对x, t方向分别进行均匀网格剖分，并标记网格格点为$x_{i+1} = x_i + \Delta x$, $t^{n+1} = t^n + \Delta t$.

(1) 构造该方程组的迎风格式，并分析其L^2模稳定性；

(2) 当$m = 1$，即为标量方程时，试分析构造格式的色散性和耗散性。

二、对于偏微分方程初边值问题

$$\begin{cases} u_t = (1+\sin^2 x)u_{xx}, & x \in (0, 2\pi),\ t > 0, \\ u(x,0) = \cos(x), & x \in [0, 2\pi], \\ u_x(0,t) = 0,\ u(1,t) = 0, & t \geq 0, \end{cases}$$

构造按最大模相容的有限差分方法，并证明之。这里，时空区域均采用均匀剖分，且$x_j = jh,\ j = 0, 1, \ldots, N;\ h = 2\pi/N$。

1.2 数值代数（20分）

一、设有线性方程组 $Ax=b$, 其中, $A=\begin{pmatrix} -2 & 1 & 0 \\ 1 & -2 & 1 \\ 0 & 1 & -2 \end{pmatrix}$, $b=\begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}$.

(1) 求矩阵 A 的 Doolittle 分解并利用该分解去解此方程组；

(2) 写出相应的 Jacobi 迭代的迭代格式(分量形式);

(3) 证明此时 Jacobi 迭代收敛.

1.3 数值分析（20分）

一、给定如下数据

$$f(-1)=1,\ f'(-1)=1,\ f'(1)=2,\ f(2)=1,$$

证明通过上述数据点的Hermite插值函数不存在。

二、给出如下积分

$$\int_{-1}^{1}|x|e^{x}dx$$

的辛普森复化积分公式，并给出该数值积分公式关于网格h的误差表达形式。

2. 专业基础（30分）

请根据专业方向和招生导师的要求, 仅选择如下五个模块之一进行解答, 并且在此用'$\surd$'标记出您所选择的模块：

（ ）2.1 有限元方法
（ ）2.2 样条函数
（ ）2.3 小波分析
（ ）2.4 统计机器学习理论
（ ）2.5 计算机算法理论

2.1 有限元方法

一、设Ω为二维多边形区域，考虑问题：

$$\begin{cases} -\Delta u = f, & \text{in } \Omega, \\ \alpha u + \frac{\partial u}{\partial \boldsymbol{n}} = g, & \text{on } \partial\Omega, \end{cases}$$

其中α为常数，$\boldsymbol{n}$为边界$\partial\Omega$单位外法向。

(1) 写出该问题的的变分形式，并给出α, f, g需要满足的条件以保证该变分问题在H^1空间存在唯一解。

(2) 设变分问题的解$u \in H^2$时，证明u也满足原微分方程。

(3) 在正则三角网格剖分下用分片线性协调有限元对该变分问题进行逼近时，试证明H^1及L^2误差估计。

2.2 样条函数

一、给定区间$[[0,6]$的一个分割$\{1,2,4,5\}$，多项式的次数为$m=3$，内部节点的重数为$\{1,3,2,1\}$，

(1) 写出扩充分割后的节点序列，给出该扩充节点系列对应的B样条基函数的节点序列，以及这些基函数张成的线性空间的分片结构，即其分片多项式以及在节点处的连续性；

(2) 计算上述B样条基函数中满足在点1处的值为零，但是导数值不为零的所有B样条的Beizer表示；

(3) 将$(x-2)_+ + (x-4)_+^2$表示成上述B样条基函数的线性组合；

2.3 小波分析

一、如果$f(t)$的傅里叶变换$\hat{f}(\lambda)$是可导的，记导函数为$\hat{f}'(\lambda)$，且$\hat{f}(0)=\hat{f}'(0)=0$，证明：

$$\int_R f(t)dt = \int_R tf(t)dt = 0. \tag{1}$$

二、给出基小波的容许条件，并证明墨西哥帽子小波：

$$\psi(t) = \frac{1}{\sqrt{2\pi}}(1-t^2)e^{-\frac{t^2}{2}} \tag{2}$$

满足基小波的容许性条件。

2.4 统计机器学习理论

一、(10分)
考虑寻找最大化分离超平面（the maximum margin separating hyperplane）的优化问题，假设类别之间是线性可分的：

$$\min_{\mathbf{w}\in\mathbb{R}^d} \frac{1}{2}\mathbf{w}^T\mathbf{w} \quad \text{subject to} \quad y_i(\mathbf{w}^T\mathbf{x}_i + \mathbf{b}) \geq 1, i = 1, \ldots, n, \tag{3}$$

其中$\{(\mathbf{x}_i, y_i), i = 1, \ldots, n\}$ 是训练数据集。证明最大分离超平面值ρ满足如下等式：

$$\frac{1}{\rho^2} = \sum_{i=1}^{n} \alpha_i^*, \tag{4}$$

其中$\alpha_i^*, i = 1, \ldots, n$是通过求解对偶问题获得的。

二、(10分)
在对高维数据（例如图像）使用Principal component analysis（PCA）或Linear discriminant analysis（LDA）时，在实践中会出现哪些数值计算问题？我们该如何解决这些问题？

三、(10分)
(1) 什么是Expectation – maximization（EM）算法？
(2) EM 最适合解决什么样的问题？
(3) 解释应用EM算法估计高斯混合模型中参数的主要思想。

2.5 计算机算法理论

一、(10分)
赫夫曼编码可以有效压缩数据，节省大量空间。它可以根据每个字符出现的频率，构造出字符的最优二进制表示。请回答以下问题：
(1)(5分)给定n个字符的集合$C = \{c_i\}$，对每个字符，其属性$c_i.fre$表示的出现频率（未进行排序）。现在想要构造它们的最优二进制编码，请写出算法的思路及伪代码。
(2)(5分)设C为斐波那契数列的前8项（未进行归一化，即c_i依次为1, 1, 2, 3, 5, 8, 13, 21），请画出编码树，并写出此字符集合的二进制赫夫曼编码结果，并写出推广到斐波那契数列的前n项时，编码结果会如何？

二、(10分)
现有一个整数序列$\{a_i\}, i = 1, \ldots, n$，各元素符号未知，但至少有一个正数。请回答以下问题：
(1)(2分)请用分治的思想，设计时间复杂度为$O(n \log n)$的算法，求出该数列中和最大的非空连续子列的元素值的和。
(2)(3分)请设计一个非递归的，时间复杂度为$O(n)$的算法，来求出该数列中和最大的非空连续子列的元素值的和。（提示：$\{a_1, \ldots, a_j\}$的和最大子列，要么与$\{a_1, \ldots, a_{j-1}\}$的和最大子列相同，要么包含a_j）
(3)(5分)请设计一个非递归的，时间复杂度为$O(n)$的算法，来求出该数列的积最大的非空连续子列的元素值的积。

三、(10分)
(1)(5分)请设计高效算法求解二分图的最大匹配问题，描述算法思想并给出相应的示意图。
(2)(5分)简述求解图的最大流问题的Push-Relabel算法思想，从以下四个方面回答：算法核心思想，两个基本操作，算法终止条件。