Problem 1

（备注，这题没有讲明，但是从后面的讨论中可以推出这里为无向图）

假设$|V_0|=m$，并且

$$\vec{v}_i=\vec{v}_i^0, i=n-m+1,\ldots,n$$

记B为邻接矩阵，即

$$B_{ij}=\begin{cases}1 & (i,j)\in E\\ 0 & \text{其他}\end{cases}$$

将其记为如下分块形式：

$$B=\left[\begin{array}{c|c} B_1 & B_2\\ \hline B_3 & B_4\end{array}\right]\in\mathbb{R}^{n\times n}$$

其中

$$B_1\in\mathbb{R}^{(n-m)\times(n-m)}, B_2\in\mathbb{R}^{(n-m)\times m}, B_3\in\mathbb{R}^{m\times(n-m)}, B_4\in\mathbb{R}^{m\times m}$$

记$1_k\in\mathbb{R}^k$表示全1的k维列向量，假设

$$\vec{v}_i=[x_i,y_i]^T$$

那么记

$$\begin{aligned}\vec{x}&=\begin{bmatrix}x_1\\ \vdots\\ x_{n-m}\end{bmatrix},\vec{y}=\begin{bmatrix}y_1\\ \vdots\\ y_{n-m}\end{bmatrix}\\ \vec{x}'&=\begin{bmatrix}x_{n-m+1}\\ \vdots\\ x_n\end{bmatrix},\vec{y}'=\begin{bmatrix}y_{n-m+1}\\ \vdots\\ y_n\end{bmatrix}\\ V_1&=[\,\vec{x}\quad\vec{y}\,],V_2=[\,\vec{x}'\quad\vec{y}'\,]\end{aligned}$$

(a)对能量式进行化简

$$\begin{aligned}E(\vec{v}_1,\ldots,\vec{v}_n)&=\sum_{(i,j)\in E}\|\vec{v}_i-\vec{v}_j\|_2^2\\ &=\sum_{(i,j)\in E}(\vec{v}_i^T\vec{v}_i-2\vec{v}_i^T\vec{v}_j+\vec{v}_j^T\vec{v}_j)\end{aligned}$$

$\forall 1\le k\le n-m$，我们有

$$\begin{aligned}\nabla_{\vec{v}_k}\Lambda &= \sum_{(k,j)\in E}(2\vec{v}_k - 2\vec{v}_j) + \sum_{(i,k)\in E}(-2\vec{v}_i + 2\vec{v}_k) \\ &= 2\left(\sum_{(k,j)\in E}\vec{v}_k + \sum_{(i,k)\in E}\vec{v}_k - \sum_{\substack{(k,j)\in E\\1\le j\le n-m}}\vec{v}_j - \sum_{\substack{(k,j)\in E\\n-m+1\le j\le n}}\vec{v}_j - \sum_{\substack{(i,k)\in E\\1\le i\le n-m}}\vec{v}_i - \sum_{\substack{(i,k)\in E\\n-m+1\le i\le n}}\vec{v}_i\right)\end{aligned}$$

令上式为$\vec{0}$得到

$$\sum_{(k,j)\in E}\vec{v}_k + \sum_{(i,k)\in E}\vec{v}_k - \sum_{\substack{(k,j)\in E\\1\le j\le n-m}}\vec{v}_j - \sum_{\substack{(i,k)\in E\\1\le i\le n-m}}\vec{v}_i - \sum_{\substack{(k,j)\in E\\n-m+1\le j\le n}}\vec{v}_j - \sum_{\substack{(i,k)\in E\\n-m+1\le i\le n}}\vec{v}_i = \vec{0}$$

写成矩阵形式为

$$\left(\begin{bmatrix}B_1 & B_2\end{bmatrix} + \begin{bmatrix}B_1\\B_3\end{bmatrix}^T\right)1_n V_1 - (B_1 + B_1^T)V_1 = (B_2 + B_3^T)V_2$$
$$\left(\operatorname{diag}\left(\left(\begin{bmatrix}B_1 & B_2\end{bmatrix} + \begin{bmatrix}B_1\\B_3\end{bmatrix}^T\right)1_n\right) - (B_1 + B_1^T)\right)V_1 = (B_2 + B_3^T)V_2$$

记

$$\begin{aligned}A &= \operatorname{diag}\left(\left(\begin{bmatrix}B_1 & B_2\end{bmatrix} + \begin{bmatrix}B_1\\B_3\end{bmatrix}^T\right)1_n\right) - (B_1 + B_1^T) \\ \vec{b}_x &= (B_2 + B_3^T)\vec{x}' \\ \vec{b}_y &= (B_2 + B_3^T)\vec{y}'\end{aligned}$$

那么将上式写为分量形式得到

$$\begin{aligned}A\vec{x} &= \vec{b}_x \\ A\vec{y} &= \vec{b}_y\end{aligned}$$

最后验证A为对称正定矩阵，对称性显然，下证正定性，首先记

$$C_1 = \begin{bmatrix}B_1 & B_2\end{bmatrix}, C_2 = \begin{bmatrix}B_1\\B_3\end{bmatrix}$$

$\forall\vec{y} = (y_1, \ldots, y_{n-m})^T$，计算$\vec{y}^T A\vec{y}$:

$$\begin{aligned}\vec{y}^T A\vec{y} &= \vec{y}^T\left(\operatorname{diag}\left(\left(C_1 + C_2^T\right)1_n\right) - (B_1 + B_1^T)\right)\vec{y} \\ &= \vec{y}^T\left(C_1 + C_2^T\right)1_n\vec{y} - \vec{y}^T(B_1 + B_1^T)\vec{y} \\ &= \sum_{i=1}^{n-m}\sum_{j=1}^{n}(B_{ij} + B_{ji})y_i^2 - \sum_{i=1}^{n-m}\sum_{j=1}^{n-m}(B_{ij} + B_{ji})y_i y_j\end{aligned}$$

因为$B_{ij}\in\{0,1\}$，所以

$$
\begin{aligned}
\sum_{i=1}^{n-m}\sum_{j=1}^{n-m}(B_{ij}+B_{ji})y_iy_j &\le \frac{1}{2}\sum_{i=1}^{n-m}\sum_{j=1}^{n-m}(B_{ij}+B_{ji})(y_i^2+y_j^2)\\
&=\sum_{i=1}^{n-m}\sum_{j=1}^{n-m}(B_{ij}+B_{ji})y_i^2
\end{aligned}
$$

因此

$$
\begin{aligned}
\vec{y}^T A\vec{y} &= \sum_{i=1}^{n-m}\sum_{j=1}^{n}(B_{ij}+B_{ji})y_i^2-\sum_{i=1}^{n-m}\sum_{j=1}^{n-m}(B_{ij}+B_{ji})y_iy_j\\
&\ge \sum_{i=1}^{n-m}\sum_{j=1}^{n}(B_{ij}+B_{ji})y_i^2-\sum_{i=1}^{n-m}\sum_{j=1}^{n-m}(B_{ij}+B_{ji})y_i^2\\
&=\sum_{i=1}^{n-m}\sum_{j=n-m+1}^{n}(B_{ij}+B_{ji})y_i^2\\
&\ge 0
\end{aligned}
$$

当且仅当$y_i=0$时等号成立，所以A对称正定。

(b)(i)将之前讨论的部分实现即可，代码如下

```
B = zeros(totalVertices);
[m, n] = size(edges);
for i = 1:m
    x = edges(i, 1);
    y = edges(i, 2);
    B(x, y) = 1;
    B(y, x) = 1;
end

B1 = B(unconstrainedVertices, unconstrainedVertices);
B2 = B(unconstrainedVertices, constrainedVertices);
B3 = B2';
B4 = B(constrainedVertices, constrainedVertices);

A = sparse(diag(([B1, B2] + [B1; B3]') * ones(totalVertices, 1)) - B1 - B1');
rhs = (B2 + B3') * constraints;
```

(ii)算法如下：

$$
\begin{aligned}
\vec{d}_k &= \vec{b}-A\vec{x}_{k-1}\\
\alpha_k &= \frac{\vec{d}_k^T\vec{d}_k}{\vec{d}_k^T A\vec{d}_k}\\
\vec{x}_k &= \vec{x}_{k-1}+\alpha_k\vec{d}_k
\end{aligned}
$$

所以对应代码如下：

```
for i=1:maxIters
    % TODO:  Update curX
    d = rhs - A * curX;
    a1 = sum(d .* d, 1);
    a2 = diag(d' * A * d)';

    alpha = a1 / a2;
    curX = curX + alpha * d;


    % Display the current iterate
    curResult(unconstrainedVertices,:) = curX;
    plotGraph(curResult, edges, f);
    title(sprintf('Gradient descent iteration %d',i));
    drawnow;
    pause(.1);
end
```

(iii)算法如下：

$$
\begin{aligned}
\text{Update search direction: } \vec{v}_k &= \vec{r}_{k-1} - \frac{\langle \vec{r}_{k-1}, \vec{v}_{k-1} \rangle_A}{\langle \vec{v}_{k-1}, \vec{v}_{k-1} \rangle_A} \vec{v}_{k-1} \\
\text{Line search: } \alpha_k &= \frac{\vec{v}_k^\top \vec{r}_{k-1}}{\vec{v}_k^\top A \vec{v}_k} \\
\text{Update estimate: } \vec{x}_k &= \vec{x}_{k-1} + \alpha_k \vec{v}_k \\
\text{Update residual: } \vec{r}_k &= \vec{r}_{k-1} - \alpha_k A \vec{v}_k
\end{aligned}
$$

代码如下：

```
%初始化
r = rhs - A * curX;
v = zeros(size(r)) + 1e-3;

for i=1:maxIters
    % TODO:  Update curX
    r1 = diag(r' * A * v)';
    v1 = diag(v' * A * v)';
    v = r - r1 ./ v1 .* v;
    alpha = diag(v' * r) ./ diag(v' * A * v);
    curX = curX + alpha' .* v;
    r = r - alpha' .* (A * v);

    % Display the current iterate
    curResult(unconstrainedVertices,:) = curX;
    plotGraph(curResult, edges, f);
    title(sprintf('Conjugate gradients iteration %d',i));
    drawnow;
    pause(.1);
end
```

(iv)共轭梯度法很快就收敛了。

Problem 2

(a)定义

$$
\begin{aligned}
\vec{x}' &= A^{-1}\vec{b} && (1)\\
\vec{x}_k &= M^{-1}\left(N\vec{x}_{k-1}+\vec{b}\right) && (2)\\
\vec{e}_k &= \vec{x}_k-\vec{x}' && (3)
\end{aligned}
$$

下面考虑$\vec{e}_k$和$\vec{e}_{k-1}$的递推关系：

$$
\begin{aligned}
\vec{e}_k &= \vec{x}_k-\vec{x}'\\
&= M^{-1}\left(N\vec{x}_{k-1}+\vec{b}\right)-\vec{x}'\\
&= M^{-1}N\left(\vec{x}_{k-1}+N^{-1}(\vec{b}-M\vec{x}')\right)
\end{aligned}
$$

注意

$$
A = M-N
$$

所以由(1)可得

$$
\begin{aligned}
A\vec{x}' = (M-N)\vec{x}' &= \vec{b}\\
\vec{b}-M\vec{x}' &= -N\vec{x}'\\
N^{-1}(\vec{b}-M\vec{x}') &= -\vec{x}'
\end{aligned}
$$

因此

$$
\begin{aligned}
\vec{e}_k &= M^{-1}N\left(\vec{x}_{k-1}+N^{-1}(\vec{b}-M\vec{x}')\right)\\
&= G(\vec{x}_{k-1}-\vec{x}')\\
&= G\vec{e}_{k-1}
\end{aligned}
$$

递推可得

$$
\vec{e}_k = G^k\vec{e}_0
$$

假设G的特征值为$\lambda_1,\ldots,\lambda_n$，对应的特征向量为$\vec{v}_1,\ldots,\vec{v}_n$，记

$$
V = \left(\,\vec{v}_1,\ldots,\vec{v}_n\,\right), \Lambda = \mathrm{diag}\{\lambda_1,\ldots,\lambda_n\}
$$

题目假设$\vec{v}_1,\ldots,\vec{v}_n$张成$\mathbb{R}^n$，那么

$$
\begin{aligned}
G &= V\Lambda V^{-1}\\
G^k &= V\Lambda^k V^{-1}
\end{aligned}
$$

因为$\lambda_i<1$，所以

$$
\begin{aligned}
\Lambda^k &\to 0 \\
G^k = V\Lambda^k V^{-1} &\to 0 \\
\vec{e}_k &\to 0
\end{aligned}
$$

因此

$$
\vec{x}_k \to \vec{x}'
$$

(b)利用圆盘定理即可：

圆盘定理

假设$A \in \mathbb{R}^{n\times n}$为$n$阶矩阵，令

$$
R_i = \sum_{j\neq i}^{n} |a_{ij}| = |a_{i1}| + \cdots + |a_{i,i-1}| + |a_{i,i+1}| + \cdots + |a_{in}|
$$

那么A的特征值z在如下圆盘中

$$
|z - a_{ii}| \leqslant R_i, i = 1, 2, \cdots, n
$$

证明：任取A的特征值λ_0，对应的特征向量为$\vec{x}$，那么

$$
A\vec{x} = \lambda_0 \vec{x}
$$

写成方程形式为

$$
\begin{cases}
a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = \lambda_0 x_1 \\
a_{21}x_1 + a_{2z}x_2 + \cdots + a_{2n}x_n = \lambda_0 x_2 \\
\cdots\cdots\cdots \\
a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nn}x_n = \lambda_0 x_n
\end{cases}
$$

设

$$
|x_r| = \max\{|x_1|, |x_2|, \cdots, |x_n|\}
$$

那么

$$
(\lambda_0 - a_{rr})\, x_r = a_{r1}x_1 + \cdots + a_{r,r-1}x_{r-1} + a_{r,r+1}x_{r+1} + \cdots + a_{rn}x_n
$$

于是

$$
\begin{aligned}
|\lambda_0 - a_{rr}|\,|x_r| \leq |a_{r1}|\,|x_1| + \cdots + |a_{r,r-1}|\,|x_{r-1}| + |a_{r,r-1}|\,|x_{r+1}| + \cdots + |a_{rn}|\,|x_n| \\
= (|a_{r1}| + \cdots + |a_{r,r-1}| + |a_{r,r\to1}| + \cdots + |a_{rn}|)\,|x_r|
\end{aligned}
$$

即

$$
|\lambda_0 - a_{rr}|\,|x_r| \leq R_r\,|x_r|
$$

但是显然$|x_r| > 0$，所以

$$
|\lambda_0 - a_{rr}| \leq R_r
$$

回到原题，我们有

$$
\begin{aligned}
R_i &= \sum_{j\neq 1}^{n} |g_{ij}| \\
&= |g_{i1}| + \cdots + |g_{i,i-1}| + |g_{i,i+1}| + \cdots + |g_{in}| \\
&= \frac{1}{|a_{ii}|}(|a_{i1}| + \cdots + |a_{i,i-1}| + |a_{i,i+1}| + \cdots + |a_{in}|) \\
&< 1
\end{aligned}
$$

而$g_{ii} = 0$，所以G的特征值满足

$$|\lambda| < R_i < 1$$

因此收敛。

Problem 3

如果

$$\sum_{i=1}^{n} a_i \vec{x}_i = \vec{0}$$

左乘$\vec{x}_k^T A, k = 1, \ldots, n$可得

$$\sum_{i=1}^{n} a_i \vec{x}_k^T A \vec{x}_i = a_k \vec{x}_k^T A \vec{x}_k + \sum_{i\neq k} a_i \vec{x}_k^T A \vec{x}_i = 0 \tag{1}$$

由A正交的定义可得

$$\vec{x}_i^T A \vec{x}_j = 0, i \neq j$$

所以(1)即为

$$a_k(\vec{x}_k^T A \vec{x}_k) = 0$$

如果A正定，$\vec{x}_k$非零，那么

$$a_k = 0, k = 1, \ldots, n$$

此时$\vec{x}_i$线性无关。

如果A半正定，那么因为$\vec{x}_k^T A \vec{x}_k$可能为0，所以无法判断a_k，即此时无法推出$\vec{x}_i$线性无关。