1.

假设函数f是满足$mI \leq \nabla^2 f(x) \leq MI$的强凸函数。令$d$为$x$处的下降方向。证明对于$0 < t \leq -\frac{\nabla f(x)^\top d}{M\|d\|_2^2}$回溯终止条件能够满足。

解:

能够满足回溯终止条件，即满足$f(x+td) \leq f(x) + \alpha t \nabla f(x)^T d$,有

$$f(x+td) = f(x) + t\nabla f(x)^T d + \frac{1}{2} t d^T \nabla^2 f(x+\theta t d) t d, \theta \in (0,1)$$

由强凸性，$\nabla^2 f(x+\theta t d) \leq MI$，代入得

$$f(x+td) \leq f(x) + \nabla f(x)^T t d + \frac{Mt^2}{2} d^T d$$

对于$0 < t \leq -\frac{\nabla f(x)^\top d}{M\|d\|_2^2}$，移项可得

$$-\frac{Mtd^T d}{2\nabla f(x)^T d} \leq \frac{1}{2}$$

又$0 < \alpha < 0.5$，则有

$$-\frac{Mtd^T d}{2\nabla f(x)^T d} \leq \frac{1}{2} \leq 1 - \alpha$$

推得$\nabla f(x)^T t d + \frac{Mt^2}{2} d^T d \leq \alpha t \nabla f(x)^T d$,

代入得

$$f(x+td) \leq f(x) + \nabla f(x)^T t d + \frac{Mt^2}{2} d^T d \leq f(x) + \alpha t \nabla f(x)^T d$$

即 $f(x+td) \leq f(x) + \alpha t \nabla f(x)^T d$，回溯终止条件能够满足得证。

2. 编程证明以上结论:

代码:

```
#include <iostream>
#include <cmath>
#include <algorithm>
#include <utility>
using namespace std;

const long double eps = 5e-13;
const long limit = 10000;
long double gamma = 0.01;

pair<long double, long double> x;

pair<long double, long double> delta(pair<long double, long double> x)
{
    pair<long double, long double> temp;
    temp.first = x.first;
    temp.second = x.second * gamma;
    return temp;
}
long double solve_fan(pair<long double, long double> x)
{
    long double sqr_sum = x.first * x.first + x.second * x.second;
    return sqrtl(sqr_sum);
}
```

```
int main()
{
    long k_cnt = 0;

    for (k_cnt = 0; k_cnt < limit; ++k_cnt)
    {

        x.first = gamma * pow((gamma - 1) / (gamma + 1), k_cnt);
        x.second = pow((1 - gamma) / (gamma + 1), k_cnt);

        if (solve_fan(delta(x)) < eps)
        {
            break;
        }

    }
    cout << "当gamma = " << gamma << " 时：" << endl;
    cout << "迭代次数k_cnt = " << k_cnt << endl;
    cout << "x = [ " << x.first << " , " << x.second << " ];" << endl;
    return 0;
}
```

结果：

γ	k
1	1
3	43
0.333333	40
100	1664
0.01	1204

由实验结果，则证明结论。

3. 对优化问题$\min\ f(x)$，迭代求解最优性条件$\nabla f(x^*)=0$。假设：

- 存在m,M满足$\forall x\in S, mI\le\nabla^2 f(x)\le MI$
- 存在常数$\gamma,\tilde\gamma\in(0,1], \|x\|\ge\gamma\|x\|_2, \|x\|_*\ge\tilde\gamma\|x\|_2$

证明：

1. 精确直线搜索时，非规范化最速下降方法的收敛性；

2. 回溯直线搜索时，非规范化最速下降方法的收敛性；

1.
存在常数 $m>0$，使得：

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

有：

$$f(x+td)=f(x)+\nabla f(x)^Ttd+\frac{1}{2}td^T\nabla^2 f(x+\theta td)td\le f(x)+\nabla f(x)^Ttd+\frac{Mt^2}{2}d^Td,\theta\in(0,1),$$

下降方向有：

$$\nabla f(x)^Td=\|\nabla f(x)\|_*\cdot\nabla f(x)^T\cdot\arg\min\{\nabla f(x)^Td:\|d\|=1\}=-\|\nabla f(x)\|_*^2.$$

代入，得：

$$f(x+td) \le f(x) - t\|\nabla f(x)\|_*^2 + \frac{Mt^2}{2\gamma^2}\|\nabla f(x)\|_*^2 \le f(x) - \frac{2\gamma^2 - Mt}{2\gamma^2}t\|\nabla f(x)\|_*^2.$$

得改进量下界:

$$f(x) - f(x+td) \ge \frac{\gamma^2}{2M}\|\nabla f(x)\|_*^2 \ge \frac{\gamma^2\tilde{\gamma}^2}{2M}\|\nabla f(x)\|_2^2.$$

由此，得:

$$(1 - \frac{\gamma^2\tilde{\gamma}^2 m}{M})(f(x) - p^*) \ge f(x+td) - p^*.$$

设定 $c = (1 - \frac{\gamma^2\tilde{\gamma}^2 m}{M}) < 1$，可以得到收敛性条件:

$$K \ge \log_{\frac{1}{c}}\left(\frac{f(x^0) - p^*}{\epsilon}\right),$$

即每次迭代的次数 K 满足上述条件，从而证明了收敛性。

2.
由回溯直线搜索停止条件,得

$$f(x+td) \le f(x) + \alpha t \nabla f(x)^T d$$

同（1）易得

$$f(x+td) \le f(x) - \frac{2\gamma^2 - Mt}{2\gamma^2}t\|\nabla f(x)\|_*^2$$

代入$t = \frac{\gamma^2}{M}$可得

$$f(x + \frac{\gamma^2}{M}d) \le f(x) + \frac{\gamma^2}{2M}\nabla f(x)^T d$$

由α范围，得

$$f(x + \frac{\gamma^2}{M}d) \le f(x) + \frac{\alpha\gamma^2}{M}\nabla f(x)^T d$$

综上，得

$$f(x+td) \le f(x) - \alpha\tilde{\gamma}^2 \min\{1, \frac{\beta\gamma^2}{M}\}\|\nabla f(x)\|_2^2$$

设定 $c = (1 - 2\alpha\tilde{\gamma}^2 m \min\{1, \frac{\beta\gamma^2}{M}\}) < 1$，可以得到收敛性条件:

$$K \ge \log_{\frac{1}{c}}(\frac{f(x^0) - p^*}{\epsilon}),$$

即每次迭代的次数 K 满足上述条件，从而证明了收敛性。

4. 对优化问题$\min\ \ f(x)$，迭代求解最优性条件$\nabla f(x^*) = 0$。假设:

- 存在m, M满足$\forall x \in S, mI \le \nabla^2 f(x) \le MI$
- 存在常数$L, \|\nabla^2 f(y) - \nabla^2 f(x)\|_2 \le L\|y - x\|_2$

证明：回溯直线搜索时牛顿方法的收敛性。

解:

阻尼牛顿阶段

存在满足$0 < \eta \leq \frac{m^2}{L}, \|\nabla f(x^k)\|_2 \geq \eta$

$$f(x^{k+1}) - f(x^k) \leq -\gamma$$

$$f(x^k + td_{nt}^k) \leq f(x^k) + t\nabla f(x^k)^T d_{nt}^k + \frac{M\|d_{nt}^k\|_2^2}{2}t^2$$

又有

$$(d_{nt}^k)^T \nabla^2 f(x^k) d_{nt}^k \geq m\|d_{nt}^k\|_2^2$$

得

$$f(x^k + td_{nt}^k) \leq f(x^k) - t\lambda(x^k)^2 + \frac{M}{2m}t^2\lambda(x^k)^2$$

$t_{min} = m/M, \alpha < 1/2$,有

$$f(x^k + t_{min}d_{nt}^k) \leq f(x^k) - \frac{m}{2M}\lambda(x^k)^2 \leq f(x^k) - \alpha\frac{m}{M}\lambda(x^k)^2$$

当$t \geq \beta t_{min} = \beta\frac{m}{M}$时

$$f(x^{k+1}) - f(x^k) \leq -\alpha\beta\frac{m}{M}\lambda(x^k)^2\ leq - \alpha\beta\frac{m}{M}\|\nabla^2 f(x^k)\|_2^2 \leq -\alpha\beta\eta^2\frac{m}{M}$$

此时易得阻尼牛顿阶段迭代次数小于等于$\frac{f(x^0)-p^*}{\gamma}$,$\gamma = \alpha\beta\eta^2\frac{m}{M}, 0 < \eta \leq \frac{m^2}{L}$

二次收敛阶段

$$\|\nabla^2 f(x^k + td) - \nabla^2 f(x^k)\|_2 \leq tL\|d\|_2^3$$

令$\tilde{f}(t) = f(x^k + td)$
推得

$$\tilde{f}(t) \leq \tilde{f}(0) - t\lambda(x^k)^2 + \frac{1}{2}t^2\lambda(x^k)^2 + \frac{t^3L}{6m^{3/2}}\lambda(x^k)^3$$

$t = 1$,代入得

$$f(x^k + d) \leq f(x^k) - \lambda(x^k)^2(\frac{1}{2} - \frac{L\lambda(x^k)}{6m^{3/2}}) \leq f(x^k) - \alpha\lambda(x^k)^2 = f(x^k) + \alpha\nabla f(x^k)^T d$$

$$f(x^{k+1}) = f(x^k + d)$$

即$\|\nabla f(x^k)\|_2 < \eta \leq 3(1-2\alpha)\frac{m^2}{L}$,则回溯直线搜索产生$t^k = 1$

$$\nabla f(x^{k+1}) = \nabla f(x^k + d) = \nabla f(x^k + d) - \nabla f(x^k) - \nabla^2 f(x^k)d$$

$$\|\nabla f(x^{k+1}\|_2 = \|\int_0^1 (\nabla^2 f(x^k + td) - \nabla^2 f(x^k))ddt\|_2 \leq \int_0^1 Lt\|d\|_2^2 dt = \frac{L}{2}\|d\|_2^2 \leq \frac{L}{2m^2}\|\nabla f(x^k)\|_2^2$$

若存在某个k，使得$\|\nabla f(x^k)\|_2 < \eta$,任意$K > k$,有

$$\|\nabla f(x^K)\|_2 < \eta$$

回溯直线搜索产生$t^K=1$,并且

$$\frac{L}{2m^2}\|\nabla f(x^{K+1}\|_2 \le (\frac{L}{2m^2}\|\nabla f(x^K)\|_2)^2$$

得

$$f(x^K)-p^* \le \frac{1}{2m}\nabla f(x^K)\|_2^2 \le \frac{2m^3}{L^2}(\frac{1}{2})^{2^{K-k+1}}$$

记$\epsilon_0=\frac{2m^3}{L^2}, K_2=K-k+1$
得

$$K_2 \ge log_2(log_2(\epsilon_0/\epsilon))$$

综上，收敛性得证。