刘程华 2018011687 计91

Prob 1. Prove that Gaussian elimination is a special case of conjugate direction method.

When Gaussian elimination is employed on a positive definite coefficient matrix $A \in R^{n\times n}$ then prove that at stage $k, A^{(k)}$ is obtained by a conjugate Gram-Schmidt algorithm using $\{e_1, \cdots, e_n\}$ where e_i is the i th natural coordinate axis in $\mathbb{R}^{n-k+1}$.

when $n = 2$

$$d_0 = e_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix} - \frac{e_1^T A e_2}{e_1^T A e_1} = -\frac{a_{12}}{a_{11}} \triangleq l$$

$$d_1 = e_2 - \frac{e_1^T A e_2}{e_1^T A e_1} e_1 = \begin{pmatrix} 0 \\ 1 \end{pmatrix} + l \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} l \\ 1 \end{pmatrix}$$

$$x^* = \frac{d_0^T b}{d_0^T A e_1} d_0 + \frac{d_1^T b}{d_1^T A e_2} d_1 = \frac{b_1}{a_{11}} \begin{pmatrix} 1 \\ 0 \end{pmatrix} + \frac{l b_1 + b_2}{l a_{21} + a_{22}} \begin{pmatrix} l \\ 1 \end{pmatrix} = \begin{pmatrix} \frac{b_1}{a_{11}} + \frac{l b_1 + b_2}{l a_{21} + a_{22}} l \\ \frac{l b_1 + b_2}{l a_{21} + a_{22}} \end{pmatrix}$$

Gaussian elimination:

$$-\frac{a_{21}}{a_{11}} \triangleq l$$

$$\begin{pmatrix} a_{11} & a_{12} & b_1 \\ a_{21} & a_{22} & b_2 \end{pmatrix} \Rightarrow \begin{pmatrix} a_{11} & a_{12} & b_1 \\ 0 & a_{22} + l a_{12} & b_2 + l b_1 \end{pmatrix}$$

$$\begin{aligned} x_2 &= \frac{l b_1 + b_2}{l a_{12} + a_{22}} \\ x_1 &= \frac{b_1 - a_{12} x_2}{a_{11}} = \frac{b_1}{a_{11}} + l x_2 \end{aligned} \Rightarrow x = \begin{pmatrix} \frac{b_1}{a_{11}} + \frac{l b_1 + b_2}{l a_{21} + a_{22}} l \\ \frac{l b_1 + b_2}{l a_{21} + a_{22}} \end{pmatrix}$$

when $n > 2$

$$\begin{aligned} d_i^\top A d_i &= d_i^\top A v_i \\ d_k = v_k - \sum_{i=0}^{k-1} \beta_{ki} d_i &= v_k - \sum_{i=0}^{k-1} \frac{d_i^\top A v_k}{d_i^\top A v_i} d_i \\ \Rightarrow \ x^* = \sum_{i=0}^{n-1} \alpha_i d_i &= \sum_{i=0}^{n-1} \frac{d_i^\top b}{d_i^\top A d_i} d_i \end{aligned}$$

Assuming n=k holds

$$\begin{aligned} A_i^{(2)} &= e_i^\top A^{(1)} - \frac{e_i^\top A^{(1)} e_i}{e_1^\top A^{(1)} e_1} \\ D^\top A D = I \Rightarrow \quad A &= \left(D^\top\right)^{-1} D^{-1} \Rightarrow A^{(2)} = T A^{(1)} \end{aligned}$$

$A^{(2)}$ looks like $\begin{bmatrix} \alpha & w^\top \\ 0 & \beta \end{bmatrix}$

$TA^{(1)}T^\top$ is Positive definite matrix.

Gaussian elimination:

$$
\begin{aligned}
\alpha_k = v_k - \sum_{i=0}^{k-1} \beta_{ki} d_i = V_k - \sum_{i=0}^{k-1} \frac{d_i^\top A v_k}{d_i^\top A v_i} d_i \\
\Rightarrow x = \sum_{i=0}^{n-1} \frac{d_i^\top b}{di^\top A d_i} d_i
\end{aligned}
$$

Prob 2. Apply the Newton's method, gradient descent, conjugate gradient to the Rosenbrock function $f\left(x_1, x_2\right) = \left(1 - x_1\right)^2 + 100\left(x_2 - x_1^2\right)^2$ .Compare and discuss the convergence rates of these methods.

（1） calculate

<1> the Newton's method

```
#library('rgl')
library('graphics')

Rose = function(x){
    (1-x[1])^2 + 100 * (x[2]-x[1]^2)^2
}
Gradient1 = function(x){
  return(matrix(c(2*(x[1]-1)-400*(x[2]-x[1]^2)*x[1], 200*(x[2]-x[1]^2)),ncol=1))
}
library(numDeriv)
Newton = function(f, startpoint_x1=1, startpoint_x2=1, IterCountMax, h=0.01, tol=1e-12){
  IterCount = 0
  epsilon = tol
  x = c()
  x0 = matrix(c(startpoint_x1,startpoint_x2),ncol=1)
  x = cbind(x,x0)
  y_value = c()
  y_value = cbind(y_value, Gradient1(x[,1]))
  while(IterCount <= IterCountMax & norm(y_value[,IterCount+1],'2') > epsilon){
    x = cbind(x, x[,IterCount+1]-solve(hessian(f,x[,IterCount+1]))%*%Gradient1(x[,IterCount+1]))
    y_value = cbind(y_value, Gradient1(x[,IterCount+1]))
    IterCount = IterCount + 1
```

```
  }
  return(data.frame('IterNum'=IterCount, 'root x'=x[1,IterCount], 'root y'=x[2,IterCount]))
}

Newton(Rose, startpoint_x1=10, startpoint_x2=10, IterCountMax=10000, h=0.01, tol=1e-12)

```

```
  IterNum root.x root.y
1       6      1      1
```

<2> gradient descent

```
Gradient_descent_Backtracking_line = function(f, startpoint_x1=1, startpoint_x2=1, IterCountMax, h=0.01, tol=1e-12){
  beta = 0.001
  t = 0.1
  epsilon = tol
  # Initial value
  Alpha2 = 0.3
  IterCount = 0
  x0 = matrix(c(startpoint_x1,startpoint_x2),ncol=1)
  x = c()
  x = cbind(x,x0)
  x = cbind(x, x[,IterCount+1]-Alpha2 * grad(f,matrix(x[,IterCount+1],ncol=1)))
  # Iteration
  while(norm(matrix(x[,IterCount+2]-x[,IterCount+1], ncol=1),'2') > epsilon & IterCount<IterCountMax){
    if(f(x[,IterCount+2] - Alpha2*grad(f,matrix(x[,IterCount+2],ncol=1))) > f(x[,IterCount+2]) - Alpha2 * t * (norm(grad(f,matrix(x[,IterCount+2],ncol=1)),'2'))^2){
      Alpha2 = beta * Alpha2
    }
    x = cbind(x, x[,IterCount+2]- Alpha2 * grad(f,matrix(x[,IterCount+2],ncol=1)))
    IterCount = IterCount + 1
  }
  return(data.frame('IterNum'=IterCount+2, 'Alpha2'=Alpha2, 'x'=x[1,IterCount+2], 'y'=x[2,IterCount+2]))
}
Gradient_descent_Backtracking_line(f_Rosen, startpoint_x1=0.01, startpoint_x2=0.01, IterCountMax=100000, h=0.01, tol=1e-12)
```

```
  IterNum  Alpha2         x         y
1  100002   3e-04 0.9999982 0.9999964
```

<3> conjugate gradient

```
> optim(par=c(0,0), fn=f_Rosen, method='CG')
$par
[1] 0.7344980 0.5382608

$value
[1] 0.07030794

$counts
function gradient
     373      101

$convergence
[1] 1

$message
NULL
```

It can be found that the Newton's method has the fastest convergence rate, the conjugate gradient method is slightly slower, and the gradient descent method is the slowest and it is difficult to converge to the correct result. The following will theoretically analyze the convergence rate of the three methods.

---

(2) the convergence rate

<1> the Newton's method

For each iteration, the iterative update rule :

$$x_{k+1} = x_k - \frac{f(x)}{f'(x)}$$

For minimizing a function, the iterative update rule :

$$x_{k+1} = x_k - \left[\nabla^2 f(x_k)\right]^{-1} \nabla f(x_k)$$

Recall that the goal of Newton method is to find the optimal $x^\star$ such that $f(x^\star) = 0$, where $f : \mathbb{R} \to \mathbb{R}$. This section shows the quadratic convergence rate of Newton method.

Assume f has continuous second derivative at $x^\star$. By the 2 nd-order Taylor approximation, we know for a ξ_k in between x_k and $x^\star$:

$$0 = f(x^\star) = f(x_k) + \underbrace{\nabla f(x_k)}_{\text{grad. at } x_k}(x^\star - x_k) + \frac{1}{2}\underbrace{\nabla^2 f(\xi_k)}_{\text{Hess. at } \xi_k}(x^\star - x_k)^2$$

Suppose $\left[\nabla^{-1} f(x_k)\right]$ exists and multiply $\left[\nabla^{-1} f(x_k)\right]$ to both sides of equation, we have

$$
\begin{aligned}
& 0=\left[\nabla^{-1} f\left(x_k\right)\right] f\left(x_k\right)+\left(x^{\star}-x_k\right)+\tfrac{1}{2}\left[\nabla^{-1} f\left(x_k\right)\right] \nabla^2 f\left(\xi_k\right)\left(x^{\star}-x_k\right)^2 \\
\Rightarrow \quad & \underbrace{\left(\left[\nabla^{-1} f\left(x_k\right)\right] f\left(x_k\right)-x_k\right)}_{-x_{k+1}}+x^{\star}=-\tfrac{1}{2}\left[\nabla^{-1} f\left(x_k\right)\right] \nabla^2 f\left(\xi_k\right) \underbrace{\left(x^{\star}-x_k\right)^2}_{\epsilon_k^2} \\
\Rightarrow \quad & \epsilon_{k+1}=-\frac{\nabla^2 f\left(\xi_k\right)}{2 \nabla f\left(x_k\right)} \epsilon_k^2
\end{aligned}
$$

Let $M=\sup_{x, y} \frac{\left|\nabla^2 f(x)\right|}{|2 \nabla f(y)|}<\infty$ is a bounded quantity, then we can say Newton method has a quadratic convergence rate by showing:

$$
\left|\epsilon_{k+1}\right| \leq M \epsilon_k^2
$$

Further if we assume $\left|\epsilon_0\right|=\left|x^{\star}-x_0\right|<1$, we can say the error ϵ_k converges to 0 with quadratic rate. p=2.

<2> gradient descent

Consider the problem

$$
x^*=\arg \min_{x \in \mathbb{R}^d} f(x)
$$

and the following gradient method

$$
x^{t+1}=x^t-\alpha \nabla f\left(x^t\right)
$$

where f is L -smooth. In Theorem 2.1 we will prove sublinear convergence under the assumption that f is convex. In Theorem 2.2 we will prove linear convergence (a stronger form of convergence) under the assumption that f is μ-strongly convex.

Smoothness

A differential function f is said to be $L-$ smooth if its gradients are Lipschitz continuous, that is

$$
\|\nabla f(x)-\nabla f(y)\| \leq L\|x-y\|
$$

If f is twice differentiable then we have, by using first order expansion

$$
\nabla f(x)-\nabla f(x+\alpha d)=\int_{t=0}^\alpha \nabla^2 f(x+t d) d
$$

followed by taking the norm gives

$$
\left\|\int_{t=0}^\alpha \nabla^2 f(x+t d) d\right\|_2 \leq L \alpha\|d\|_2
$$

Dividing by α

$$
\frac{\left\|\int_{t=0}^\alpha \nabla^2 f(x+t d) d\right\|_2}{\alpha} \leq L\|d\|_2
$$

then dividing through by $\|d\|$ with $d \neq 0$ and taking the limit as $\alpha \rightarrow 0$ we have that

$$
\frac{\left\|\int_{t=0}^\alpha \nabla^2 f(x+t d) d\right\|_2}{\alpha\|d\|}=\frac{\left\|\alpha \nabla^2 f(x) d\right\|_2}{\alpha\|d\|}+O(\alpha) \underset{\alpha \rightarrow 0}{=} \frac{\left\|\alpha \nabla^2 f(x) d\right\|_2}{d} \leq L, \quad \forall d \neq 0 \in \mathbb{R}^n
$$

Taking the supremum over $d \neq 0 \in \mathbb{R}^d$ in the above gives

$$\nabla^2 f(x) \preceq LI$$

Furthermore, using the Taylor expansion of $f(x)$ and the uniform bound over Hessian we have that

$$f(y) \leq f(x) + \langle \nabla f(x), y - x \rangle + \frac{L}{2} \|y - x\|_2^2$$

Some direct consequences of the smoothness are given in the following lemma.

Lemma 2.1 If f is $L-$ smooth then

$$f\left(x - \frac{1}{L}\nabla f(x)\right) - f(x) \leq -\frac{1}{2L}\|\nabla f(x)\|_2^2$$

and

$$f\left(x^*\right) - f(x) \leq -\frac{1}{2L}\|\nabla f(x)\|_2^2$$

hold for all $x \in \mathbb{R}^d$.

Proof: The first inequality follows by inserting $y = x - \frac{1}{L}\nabla f(x)$ in the definition of smoothness since

$$\begin{aligned} f\left(x - \frac{1}{L}\nabla f(x)\right) &\leq f(x) - \frac{1}{L}\langle \nabla f(x), \nabla f(x) \rangle + \frac{L}{2}\left\|\frac{1}{L}\nabla f(x)\right\|_2^2 \\ &= f(x) - \frac{1}{2L}\|\nabla f(x)\|_2^2 \end{aligned}$$

Furthermore, by using first inequality combined with $f\left(x^*\right) \leq f(y) \quad \forall y$, we get the second inequality Indeed since

$$f\left(x^*\right) - f(x) \leq f\left(x - \frac{1}{L}\nabla f(x)\right) - f(x) \leq -\frac{1}{2L}\|\nabla f(x)\|_2^2$$

Lemma 2.2 If $f(x)$ is convex and $L-$smooth then

$$\begin{gathered} f(y) - f(x) \leq \langle \nabla f(y), y - x \rangle - \frac{1}{2L}\|\nabla f(y) - \nabla f(x)\|_2^2 \\ \langle \nabla f(y) - \nabla f(x), y - x \rangle \geq \frac{1}{L}\|\nabla f(x) - \nabla f(y)\| \quad \text{(Co-coercivity)} \end{gathered}$$

Proof: To prove this, it follows that

$$\begin{aligned} f(y) - f(x) &= f(y) - f(z) + f(z) - f(x) \\ &\leq \langle \nabla f(y), y - z \rangle + \langle \nabla f(x), z - x \rangle + \frac{L}{2}\|z - x\|_2^2 \end{aligned}$$

Minimizing in z we have that

$$z = x - \frac{1}{L}(\nabla f(x) - \nabla f(y))$$

Substituting this in gives

$$\begin{aligned}
f(y)-f(x) &= \left\langle \nabla f(y), y-x+\frac{1}{L}(\nabla f(x)-\nabla f(y))\right\rangle - \frac{1}{L}\langle \nabla f(x), \nabla f(x)-\nabla f(y)\rangle + \frac{1}{2L}\|\nabla f(x)-\nabla f(y)\|_2^2 \\
&= \langle \nabla f(y), y-x\rangle - \frac{1}{L}\|\nabla f(x)-\nabla f(y)\|_2^2 + \frac{1}{2L}\|\nabla f(x)-\nabla f(y)\|_2^2 \\
&= \langle \nabla f(y), y-x\rangle - \frac{1}{2L}\|\nabla f(x)-\nabla f(y)\|_2^2
\end{aligned}$$

Finally the second inequality follows from applying the first inequality once

$$f(y)-f(x) \leq \langle \nabla f(y), y-x\rangle - \frac{1}{2L}\|\nabla f(y)-\nabla f(x)\|_2^2$$

then interchanging the roles of x and y to get

$$f(x)-f(y) \leq \langle \nabla f(x), x-y\rangle - \frac{1}{2L}\|\nabla f(y)-\nabla f(x)\|_2^2$$

Finally adding together the two above inequalities gives

$$0 \leq \langle \nabla f(y)-\nabla f(x), y-x\rangle - \frac{1}{L}\|\nabla f(y)-\nabla f(x)\|_2^2$$

Strong convexity

We can "strengthen" the notion of convexity by defining $\mu-$ strong convexity, that is

$$f(y) \geq f(x)+\langle \nabla f(x), y-x\rangle + \frac{\mu}{2}\|y-x\|_2^2, \quad \forall x, y \in \mathbb{R}^d$$

Theorem 2.1 Let f be convex and L - smooth and let x^t for $t=1,\ldots,n$ be the sequence of iterates generated by the gradient method. It follows that

$$f\left(x^n\right)-f\left(x^*\right) \leq \frac{2L\left\|x^1-x^*\right\|^2}{n-1}$$

Proof: Let f be convex and L -smooth. It follows that

$$\begin{aligned}
\left\|x^{t+1}-x^*\right\|_2^2 &= \left\|x^t-x^*-\frac{1}{L}\nabla f\left(x^t\right)\right\|_2^2 \\
&= \left\|x^t-x^*\right\|_2^2 - 2\frac{1}{L}\left\langle x^t-x^*, \nabla f\left(x^t\right)\right\rangle + \frac{1}{L^2}\left\|\nabla f\left(x^t\right)\right\|_2^2 \\
&\leq \left\|x^t-x^*\right\|_2^2 - \frac{1}{L^2}\left\|\nabla f\left(x^t\right)\right\|_2^2
\end{aligned}$$

The inequality is because of Lemma 2.2

Thus $\left\|x^t-x^*\right\|_2^2$ is a decreasing sequence in t. Calling upon Lemma 2.1 and subtracting $f\left(x^*\right)$ from both sides gives

$$f\left(x^{t+1}\right)-f\left(x^*\right) \leq f\left(x^t\right)-f\left(x^*\right)-\frac{1}{2L}\left\|\nabla f\left(x^t\right)\right\|_2^2 \qquad (*)$$

Applying convexity we have that

$$
\begin{aligned}
f\left(x^t\right)-f\left(x^*\right) &\leq \left\langle \nabla f\left(x^t\right), x^t-x^* \right\rangle \\
&\leq \left\|\nabla f\left(x^t\right)\right\|_2 \left\|x^t-x^*\right\| \leq_{(*)} \left\|\nabla f\left(x^t\right)\right\|_2 \left\|x^1-x^*\right\|
\end{aligned}
$$

Isolating $\left\|\nabla f\left(x^t\right)\right\|_2$ in the above and inserting in $(*)$ gives

$$
f\left(x^{t+1}\right)-f\left(x^*\right) \quad \leq \underbrace{\frac{1}{2L}\frac{1}{\left\|x^1-x^*\right\|^2}}_{\beta}\left(f\left(x^t\right)-f\left(x^*\right)\right)^2
$$

Let $\delta_t=f\left(x^t\right)-f\left(x^*\right)$. since $\delta_{t+1}\leq\delta_t$ Manipulating the above we have that

$$
\delta_{t+1}\leq\delta_t-\beta\delta_t^2 \overset{\times\frac{1}{\delta_t\delta_{t+1}}}{\Leftrightarrow} \beta\frac{\delta_t}{\delta_{t+1}}\leq\frac{1}{\delta_{t+1}}-\frac{1}{\delta_t} \overset{\delta_{t+1}<\delta_t}{\Leftrightarrow} \beta\leq\frac{1}{\delta_{t+1}}-\frac{1}{\delta_t}
$$

Summing up both sides over $t=1,\ldots,n-1$ and using telescopic cancellation we have that

$$
(n-1)\beta\leq\frac{1}{\delta_n}-\frac{1}{\delta_1}\leq\frac{1}{\delta_n}
$$

Theorem 2.2 Let f be L -smooth and μ -strongly convex. From a given $x_0\in\mathbb{R}^d$ and $\frac{1}{L}\geq\alpha>0$, the iterates

$$
x^{t+1}=x^t-\alpha\nabla f\left(x^t\right)
$$

converge according to

$$
\left\|x^{t+1}-x^*\right\|_2^2\leq(1-\alpha\mu)^{t+1}\left\|x^0-x^*\right\|_2^2
$$

In particular, or $\alpha=\frac{1}{L}$ the iterates enjoy a linear convergence with a rate of μ/L. Proof: we have that

$$
\begin{aligned}
\left\|x^{t+1}-x^*\right\|_2^2 &= \left\|x^t-x^*-\alpha\nabla f\left(x^t\right)\right\|_2^2 \\
&= \left\|x^t-x^*\right\|_2^2-2\alpha\left\langle\nabla f\left(x^t\right), x^t-x^*\right\rangle+\alpha^2\left\|\nabla f\left(x^t\right)\right\|_2^2 \\
&\leq (1-\alpha\mu)\left\|x^t-x^*\right\|_2^2-2\alpha\left(f\left(x^t\right)-f\left(x^*\right)\right)+\alpha^2\left\|\nabla f\left(x^t\right)\right\|_2^2 \\
&\leq (1-\alpha\mu)\left\|x^t-x^*\right\|_2^2-2\alpha\left(f\left(x^t\right)-f\left(x^*\right)\right)+2\alpha^2L\left(f\left(x^t\right)-f\left(x^*\right)\right) \\
&= (1-\alpha\mu)\left\|x^t-x^*\right\|_2^2-2\alpha(1-\alpha L)\left(f\left(x^t\right)-f\left(x^*\right)\right)
\end{aligned}
$$

The first inequality sign is due to strong convexity, and the second inequality sign is due to Lemma 2.1.

since $\frac{1}{L}\geq\alpha$ we have that $-2\alpha(1-\alpha L)$ is negative, and thus can be safely dropped to give

$$
\left\|x^{t+1}-x^*\right\|_2^2\leq(1-\alpha\mu)\left\|x^t-x^*\right\|_2^2
$$

It now remains to unroll the recurrence.

<3> conjugate gradient

The conjugate gradient method is a method for minimizing the following quadratic functional:

$$x_* = \arg\min_{x\in\mathbb{R}^n} \varphi(x), \quad \varphi(x) = \frac{1}{2}x^T Ax - b^T x$$

Let x_0 be given initial guess.

$$\begin{aligned}
&\text{Set } r_0 = Ax_0 - b \text{ and } p_0 = -r_0, k = 0\\
&\text{While } r_k \neq 0 \text{ do :}\\
&\quad \alpha_k = \frac{\|r_k\|^2}{\|p_k\|_A^2} \quad x_{k+1} = x_k + \alpha_k p_k\\
&\quad r_{k+1} = r_k + \alpha_k A p_k\\
&\quad \beta_k = \frac{\|r_{k+1}\|^2}{\|r_k\|^2}\\
&\quad p_{k+1} = -r_k + \beta_k p_k\\
&\quad \text{Set } k = k+1\\
&\quad \text{end While}
\end{aligned}$$

The error after t iterations of the CG algorithm can be bounded as follows:

$$\|x_* - x_t\|_A \leq \frac{2}{\left(\frac{\sqrt{\kappa}+1}{\sqrt{\kappa}-1}\right)^t + \left(\frac{\sqrt{\kappa}-1}{\sqrt{\kappa}+1}\right)^t}\|x_* - x_0\|_A \leq 2\left(\frac{\sqrt{\kappa}-1}{\sqrt{\kappa}+1}\right)^t \|x_* - x_0\|_A$$

$where\ \kappa = \kappa(A) = \lambda_n/\lambda_1$ is the condition number of A.

The method is typically linear and its speed is determined by the condition number $\kappa(A)$ of the system matrix A : the larger $\kappa(A)$ is, the slower the method.

Due to words limitations, we will not give a specific proof here.