智能优化算法

第3章

凸优化问题求解算法

作业1

给定数据集$\{(x_i, y_i)\}_{i=1}^{m}$，其中$x_i=(x_{i1}, x_{i2}, \cdots, x_{in})$

请写出使用梯度下降法求解以下优化问题的步骤

$$\min_{w} h(w)=\sum_{i=1}^{m}(w_1 x_{i1}+w_2 x_{i2}+\cdots+w_n x_{in}-y_i)^2$$

$$\nabla h(w)=(\frac{\partial h}{\partial w_1},\cdots\frac{\partial h}{\partial w_l},\cdots,\frac{\partial h}{\partial w_n})$$

$$\frac{\partial h}{\partial w_l}=\sum_{i=1}^{m}(w_1 x_{i1}+w_2 x_{i2}+\cdots+w_n x_{in}-y_i)x_{il}$$

$$\underline{w_l^{k+1}=w_l^{k}}-\eta\nabla h(w^{k}),\quad l=1,2,\cdots,n$$

3.7 交替方向乘子方法

Alternating Direction Method of Multipliers

简称：**ADMM**

用于求解带等式约束的凸优化问题

下一页 上一页

3.3 ADMM 在图像恢复中的应用

$$\min_{x} \quad \frac{1}{2}\|Ax-b\|_2^2+\lambda\|x\|_2^2$$

梯度下降法求解

$$\nabla f(x)=A^T(Ax-b)+2\lambda x$$

$$\begin{aligned} x^{k+1} &= x^k-\alpha\nabla f(x^{\mathrm{k}}) \\ &= x^k-\alpha[A^T(Ax^k-b)+\lambda x^k] \end{aligned}$$

3.4 ADMM 在图像恢复中的应用

经典稀疏优化问题：lasso模型

$$\min_{x} \quad \frac{1}{2}\|Ax-b\|_2^2+\lambda\|x\|_1$$

L1范数为稀疏范数，具有良好的保边性质

梯度下降法求解

$$\nabla f(x)=A^T(Ax-b)+\lambda\frac{x}{|x|}$$

在0处不可微

求解算法不稳定

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

$$=x^k-\alpha[A^T(Ax^k-b)+\lambda\frac{x^k}{|x^k|}]$$

可通过引入新变量，分离出该不可微项

如何分离出该不可微项？？？？

$$\min_{x} \quad \frac{1}{2}\|Ax-b\|_2^2+\lambda\|x\|_1$$

$$\begin{cases} \min\limits_{x,y} & \frac{1}{2}\|Ax-b\|_2^2+\lambda\|y\|_1 \\ \text{s.t.} & x=y \end{cases}$$

转换该等式约束优化问题

增广拉格朗日函数

$$L_c(x,y,v)=\frac{1}{2}\|Ax-b\|_2^2+\lambda\|y\|_1+v^T(x-y)+\frac{c}{2}\|x-y\|_2^2$$

① $\boldsymbol{x}$ 的子问题

$$\boldsymbol{x}^{k+1}=\arg\min_{\boldsymbol{x}}\frac{1}{2}\|A\boldsymbol{x}-\boldsymbol{b}\|_2^2+\left(\boldsymbol{v}^k\right)^T\boldsymbol{x}+\frac{c}{2}\left\|\boldsymbol{x}-\boldsymbol{y}^k\right\|_2^2$$

② $\boldsymbol{y}$ 的子问题

$$\boldsymbol{y}^{k+1}=\arg\min_{\boldsymbol{y}}\lambda\|\boldsymbol{y}\|_1-\left(\boldsymbol{v}^k\right)^T\boldsymbol{y}+\frac{c}{2}\left\|\boldsymbol{y}-\boldsymbol{x}^{k+1}\right\|_2^2$$

③ $\boldsymbol{v}$ 的子问题

$$\boldsymbol{v}^{k+1}=\boldsymbol{v}^k+c\left(\boldsymbol{x}^{k+1}-\boldsymbol{y}^{k+1}\right)$$

作业2

- 请写出使用ADMM算法求解以下优化问题的算法详细步骤

$$\min_{\boldsymbol{u}} \|\nabla \boldsymbol{u}\|_1 + \alpha \|\boldsymbol{u} - \boldsymbol{u}^0\|_2^2$$

其中 $\|\nabla \boldsymbol{u}\|_1$ 有以下定义：

（1） $\sum_{ij} \sqrt{(\boldsymbol{u}_{ij})_x^2 + (\boldsymbol{u}_{ij})_y^2}$；

- 由于**L1**范数不可微，需要将第一项与第二项分离开进行求解：

求解1：引入变量p,分离L1范数

$$\min_{p,u} \| p \|_1 + \frac{\alpha}{2} \| u - u^0 \|_2^2,$$

$$\text{s.t.} \quad \mathrm{p} = \nabla \mathrm{u}$$

求解2：定义增广拉格朗日函数：

$$\mathrm{L}(\mathrm{u}, \mathrm{p}, \lambda) = \| \boldsymbol{p} \|_1 + \frac{\boldsymbol{\alpha}}{2} \| \boldsymbol{u} - \boldsymbol{u}^0 \|_2^2 + < \boldsymbol{\lambda}, \boldsymbol{p} - \nabla \boldsymbol{u} > + \frac{\boldsymbol{\beta}}{2} \| \boldsymbol{p} - \nabla \boldsymbol{u} \|_2^2$$

$$\boldsymbol{p} = (\boldsymbol{p}_1, \boldsymbol{p}_2), \mathrm{div}(\boldsymbol{p}) = \boldsymbol{p}_{1x} + \boldsymbol{p}_{2y}$$

下一页　上一页

求解3 : ADMM算法迭代求解:

u子问题:固定$\lambda^{\mathrm{k}}, p^{k}$,求解$u^{k+1}$:

$$\min_{\mathrm{u}} \frac{\alpha}{2} \| u-u^{0} \|_{2}^{2} + < \lambda^{k}, -\nabla u > + \frac{\beta}{2} \| p^{k} - \nabla u \|_{2}^{2}$$

p子问题:固定$\lambda^{\mathrm{k}}, u^{k+1}$,求解$p^{k+1}$:

$$\min_{\mathrm{p}} \| p \|_{1} + < \lambda^{k}, p > + \frac{\beta}{2} \| p - \nabla u^{k+1} \|_{2}^{2}$$

更新λ: $\lambda^{\mathrm{k+1}} = \lambda^{k} + \beta(p^{k+1} - \nabla u^{k+1})$

$$\min_{u} f(u) = \frac{\alpha}{2} \| u - u^0 \|_2^2 + < \lambda^k, -\nabla u > + \frac{\beta}{2} \| p^k - \nabla u \|_2^2$$

$$\nabla f(u) = 0;$$

$$\nabla(\frac{\alpha}{2} \| u - u^0 \|_2^2) = \alpha(u - u^0);$$

$$\nabla(< \lambda^k, -\nabla u >) = -\nabla(\mathrm{div}(\lambda^k), u) = \mathrm{div}(\lambda^k);$$

$$\nabla(\frac{\beta}{2} \| p^k - \nabla u \|_2^2) = \nabla(\frac{\beta}{2} < p^k - \nabla u, p^k - \nabla u >);$$

$$= \nabla(\frac{\beta}{2}(< p^k, \ p^k > -2 < p^k, \nabla u > + < \nabla u, \nabla u >));$$

$$= \nabla(\frac{\beta}{2}(2 < \mathrm{div}(p^k), u > -2 < \mathrm{div}(\nabla u), u >))$$

$$= \beta(\mathrm{div}(p^k) - \mathrm{div}(\nabla u))$$

$$\nabla f(u) = 0$$

$$\beta(\mathrm{div}(p^k) - \mathrm{div}(\nabla u)) + \alpha(u - u^0) + \mathrm{div}(\lambda^k) = 0$$

$$-\beta \mathrm{div}(\nabla u) + \alpha u = -\beta \mathrm{div}(p^k) - \mathrm{div}(\lambda^k) + \alpha u^0$$

$$-\beta \Delta u + \alpha I u = -\beta \mathrm{div}(p^k) - \mathrm{div}(\lambda^k) + \alpha u^0$$

$$\min_{p} \| \boldsymbol{p} \|_1 + < \lambda^k , \boldsymbol{p} > + \frac{\beta}{2} \| \boldsymbol{p} - \nabla \boldsymbol{u}^{k+1} \|_2^2$$

$$< \lambda^k , \boldsymbol{p} > + \frac{\beta}{2} \| \boldsymbol{p} - \nabla \boldsymbol{u}^{k+1} \|_2^2$$

$$= < \lambda^k , \boldsymbol{p} > + \frac{\beta}{2} < \boldsymbol{p} - \nabla \boldsymbol{u}^{k+1} , \quad \boldsymbol{p} - \nabla \boldsymbol{u}^{k+1} >$$

$$=< \lambda^k , \boldsymbol{p} > + \frac{\beta}{2} (< \boldsymbol{p}, \boldsymbol{p} > -2 < \boldsymbol{p}, \nabla \boldsymbol{u}^{k+1} > + \ < \nabla \boldsymbol{u}^{k+1}, \nabla \boldsymbol{u}^{k+1} >)$$

$$= \frac{\beta}{2} (< \boldsymbol{p}, \boldsymbol{p} > -2 < \boldsymbol{p}, \nabla \boldsymbol{u}^{k+1} - \frac{1}{\beta} \lambda^k >)$$

$$= \frac{\beta}{2} < \boldsymbol{p} - (\nabla \boldsymbol{u}^{k+1} - \frac{1}{\beta} \lambda^k), \boldsymbol{p} - (\nabla \boldsymbol{u}^{k+1} - \frac{1}{\beta} \lambda^k) >= \frac{\beta}{2} \| \boldsymbol{p} - (\nabla \boldsymbol{u}^{k+1} - \frac{1}{\beta} \lambda^k) \|_2^2$$

$$\min_{p} \| \boldsymbol{p} \|_1 + \frac{\beta}{2} \| \boldsymbol{p} - (\nabla \boldsymbol{u}^{k+1} - \frac{1}{\beta} \lambda^k) \|_2^2$$

下一页 上一页

$$\min_{p} \|\boldsymbol{p}\|_1 + \frac{\beta}{2}\|\boldsymbol{p} - (\nabla \boldsymbol{u}^{k+1} - \frac{1}{\beta}\lambda^k)\|_2^2$$

记$\nabla \boldsymbol{u}^{k+1} - \frac{1}{\beta}\lambda^k = \boldsymbol{c}$

$$\min_{p} \|\boldsymbol{p} - 0\|_1 + \frac{\beta}{2}\|\boldsymbol{p} - \boldsymbol{c}\|_2^2$$

$$\min_{p} \sum_{i,j}^{M,N} |\boldsymbol{p}_{ij}| + \frac{\beta}{2}\sum_{ij}^{M,N}(p_{ij} - c_{ij})^2$$

$$= \min_{p} \frac{\beta}{2}\sum_{ij}^{M,N}[(p_{ij} - c_{ij})^2 + |\boldsymbol{p}_{ij}|]$$

转化为对每个像素求解子问题

$$\min_{p_{ij}} \frac{\beta}{2}(\boldsymbol{p}_{ij} - c_{ij})^2 + |\boldsymbol{p}_{ij}|$$

下一页 上一页

转化为对每个像素求解子问题

$$\min_{p_{ij}} \frac{\beta}{2}(\boldsymbol{p}_{ij} - c_{ij})^2 + |\boldsymbol{p}_{ij}|$$

$$\min_{p_{ij}} \frac{\beta}{2}(\boldsymbol{p}_{ij} - c_{ij})^2 + |\boldsymbol{p}_{ij} - 0|$$

几何意义：$\boldsymbol{p}_{ij}$到原点$(0,0)$, c_{ij}距离之和最短

$\boldsymbol{p}_{ij}$,原点，$\boldsymbol{c}_{ij}$点三点共线，即满足：

$\boldsymbol{p}_{ij} = \boldsymbol{k}\boldsymbol{c}_{ij}$，$0 \le \boldsymbol{k} \le 1$，代入优化问题

$$\min_{k} |\boldsymbol{k}\boldsymbol{c}_{ij}| + \frac{\beta}{2}(\boldsymbol{k}\boldsymbol{c}_{ij} - \boldsymbol{c}_{ij})^2$$

$$\min_{k} k|c_{ij}| + \frac{\beta}{2}(k-1)^2 c_{ij}^{\ 2}$$

关于k求导数

$$|c_{ij}| + \beta(k-1)c_{ij}^{\ 2} = 0$$

$$k = 1 - \frac{1}{\beta|c_{ij}|}, \quad 0 \le k \le 1$$

$$k = 1 - \frac{1}{\beta|c_{ij}|} \ge 0$$

$$\beta|c_{ij}| \ge 1$$

转化为对每个像素求解子问题

$$\min_{p_{ij}} \frac{\beta}{2}(p_{ij} - c_{ij})^2 + |p_{ij}|$$

$$p_{ij} = \begin{cases} (1 - \frac{1}{\beta|c_{ij}|})|c_{ij}|, \beta|c_{ij}| \ge 1 \\ 0, else \end{cases}$$

$$\nabla u_{ij}^{\ k+1} - \frac{1}{\beta}\lambda_{ij}^{\ k} = c_{ij}$$

ADMM算法：

1.初始化：$\boldsymbol{p}^k = 0, \boldsymbol{\lambda}^k = 0, \boldsymbol{k} = 0$

2. do

求解u - 子问题 $\boldsymbol{u}^{k+1}$

$$-\boldsymbol{\beta}\Delta\boldsymbol{u} + \boldsymbol{\alpha I u} = -\boldsymbol{\beta}\mathrm{div}(\boldsymbol{p}^k) - \mathrm{div}(\boldsymbol{\lambda}^k) + \boldsymbol{\alpha u}^0$$

求解p - 子问题 $\boldsymbol{p}^{k+1}$

$$\boldsymbol{p}_{ij} = \begin{cases} (1 - \dfrac{1}{\boldsymbol{\beta} \mid \boldsymbol{c}_{ij} \mid}) \mid \boldsymbol{c}_{ij} \mid, \boldsymbol{\beta} \mid \boldsymbol{c}_{ij} \mid \geq 1 \\ 0, \boldsymbol{else} \end{cases}$$

更新$\boldsymbol{\lambda}$:

$$\boldsymbol{\lambda}^{\mathrm{k}+1} = \boldsymbol{\lambda}^k + \boldsymbol{\beta}(\boldsymbol{p}^{k+1} - \nabla \boldsymbol{u}^{k+1})$$

$$\boldsymbol{while}(\| \boldsymbol{u}^{k+1} - \boldsymbol{u}^k \|_2^2 \geq 0.000001)$$

程序参考

```
clear
clc
close all
%%
img1 = im2double(imread('origin.jpg'));
img1 = rgb2gray(img1);
figure, imshow(img1);
title('clean image');
[n2,m2] = size(img1);
u0 = imnoise(img1, 'gaussian',0.005);
figure, imshow(u0);
title('noisy image');
```

填充laplace矩阵

```
L = sparse(n2*m2,n2*m2);
%%%%
L(1,1) = -2;        L(1,2) = 1;             L(1,m2+1) = 1;
L(m2,m2) = -2;      L(m2,m2-1) = 1;         L(m2,2*m2) = 1;
L(m2*(n2-1)+1,m2*(n2-1)+1) = -2;        L(m2*(n2-1)+1,m2*(n2-1)+2) = 1;    L(m2*(n2-1)+1,m2*(n2-2)+1) = 1;
L(m2*n2,m2*n2) = -2;     L(m2*n2,m2*n2-1) = 1;  L(m2*n2,m2*(n2-1)) = 1;
for i=2:m2-1
    L(i,i) = -3;    L(i,i+1) = 1;     L(i,i-1) = 1;  L(i,i+m2) = 1;
    L(m2*(n2-1)+i,m2*(n2-1)+i) = -3;   L(m2*(n2-1)+i,m2*(n2-1)+i-1) = 1;
    L(m2*(n2-1)+i,m2*(n2-1)+i+1) = 1;   L(m2*(n2-1)+i,m2*(n2-2)+i) = 1;
end
for i=2:n2-1
    L(m2*(i-1)+1,m2*(i-1)+1) = -3;  L(m2*(i-1)+1,m2*(i-2)+1) = 1; L(m2*(i-1)+1,m2*(i)+1) = 1; L(m2*(i-1)+1,m2*(i-1)+2) = 1;
    L(m2*i,m2*i) = -3;    L(m2*i,m2*i-1) = 1; L(m2*i,m2*(i-1)) = 1;  L(m2*i,m2*(i+1)) = 1;
end
for i=2:n2-1
    for j=2:m2-1
        L(m2*(i-1)+j,m2*(i-1)+j) = -4;
        L(m2*(i-1)+j,m2*(i-1)+j+1) = 1;
        L(m2*(i-1)+j,m2*(i-1)+j-1) = 1;
        L(m2*(i-1)+j,m2*(i-2)+j) = 1;
        L(m2*(i-1)+j,m2*i+j) = 1;
    end
end
%%%
L = -L;
```

```
%% p subproblem

grad_x = [diff(u,1,2), u(:,1,:) - u(:,end,:)];
grad_y = [diff(u,1,1); u(1,:,:) - u(end,:,:)];
w1 = grad_x -lambda_x/r;
w2 = grad_y -lambda_y/r;
w = sqrt(w1.^2 + w2.^2);
px = (1 - 1./(r*w)).*w1;   py = (1 - 1./(r*w)).*w2;
```

```
%% update_lagrangian multiplier

lambda_x = lambda_x + r*(px-grad_x);
lambda_y = lambda_y + r*(py-grad_y);
```