Chi-Square fitting

2017135002/최성윤

기본원리

$$\sum_{i=1}^{n} \frac{[y_i - y(x_i)]^2}{\sigma_i^2}$$ 최소화

$$y(x) = y(x; a, b) = a + bx$$

$$\chi^2(a, b) = \sum_{i=1}^{N} \left(\frac{y_i - a - bx_i}{\sigma_i} \right)^2$$

Function Minimization으로

$$0 = \frac{\partial \chi^2}{\partial a} = -2 \sum_{i=1}^{N} \frac{y_i - a - bx_i}{\sigma_i^2}$$

$$0 = \frac{\partial \chi^2}{\partial b} = -2 \sum_{i=1}^{N} \frac{x_i(y_i - a - bx_i)}{\sigma_i^2}$$

$$a = \frac{S_{xx}S_y - S_x S_{xy}}{\Delta}$$

$$b = \frac{S S_{xy} - S_x S_y}{\Delta}$$

$$\Delta \equiv S S_{xx} - (S_x)^2$$

$$S \equiv \sum_{i=1}^{N} \frac{1}{\sigma_i^2} \quad S_x \equiv \sum_{i=1}^{N} \frac{x_i}{\sigma_i^2} \quad S_y \equiv \sum_{i=1}^{N} \frac{y_i}{\sigma_i^2}$$

$$S_{xx} \equiv \sum_{i=1}^{N} \frac{x_i^2}{\sigma_i^2} \quad S_{xy} \equiv \sum_{i=1}^{N} \frac{x_i y_i}{\sigma_i^2}$$

$$\sigma_a^2 = S_{xx}/\Delta$$

$$\sigma_b^2 = S/\Delta$$

실행

$$S \equiv \sum_{i=1}^{N} \frac{1}{\sigma_i^2} \quad S_x \equiv \sum_{i=1}^{N} \frac{x_i}{\sigma_i^2} \quad S_y \equiv \sum_{i=1}^{N} \frac{y_i}{\sigma_i^2}$$

$$S_{xx} \equiv \sum_{i=1}^{N} \frac{x_i^2}{\sigma_i^2} \quad S_{xy} \equiv \sum_{i=1}^{N} \frac{x_i y_i}{\sigma_i^2} \quad \Delta \equiv S S_{xx} - (S_x)^2$$

```
s=np.sum(1/yer**2)
sx=np.sum(x/(yer**2))
sy=np.sum(y/(yer**2))
sxx=np.sum((x**2)/(yer**2))
sxy=np.sum(x*y/(yer**2))
d=s*sxx-(sx**2)
```

$$a = \frac{S_{xx}S_y - S_x S_{xy}}{\Delta}$$

$$b = \frac{S S_{xy} - S_x S_y}{\Delta}$$

```
a=(sxx*sy-sx*sxy)/d
b=(s*sxy-sx*sy)/d
```

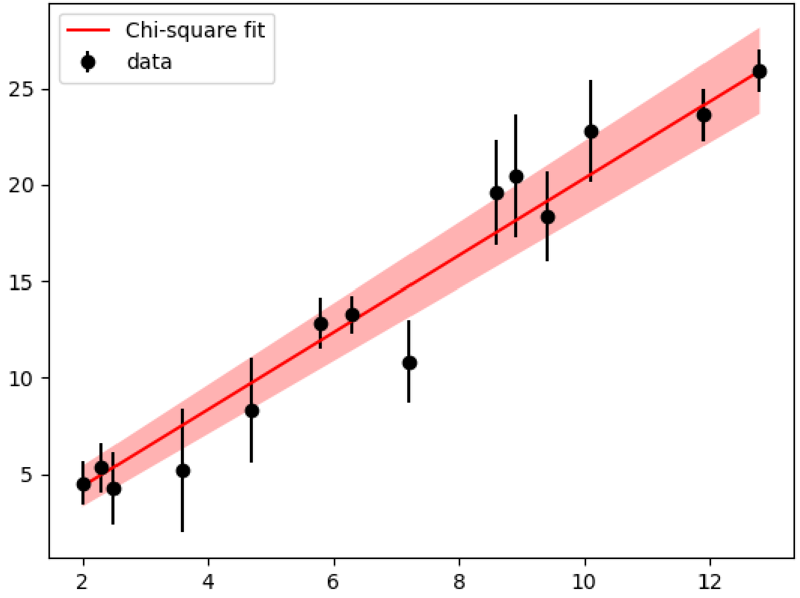
결과

a,b 값

```
0.37712438758161604
1.993951076761468З
```

a,b error값

```
0.8438718500969482
0.10806292881858451
```



토의사항

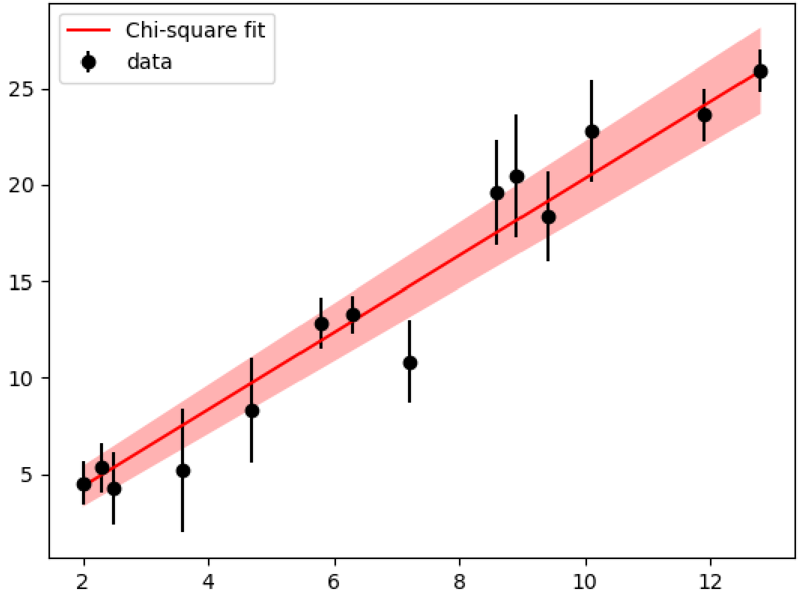
1.Reduced Chi-Square?

$$\chi^2_{reduced} = \frac{\chi^2}{M}$$ M=14-12=2

Reduced chi square 값

```
0.6001758574841355
```

적절하게 fitting 되었다



2. y뿐만 아니라 x에도 오차가 있는 경우?

$$\sum_{i=1}^{n} \frac{[y_i - y(x_i)]^2}{\sigma_i^2}$$

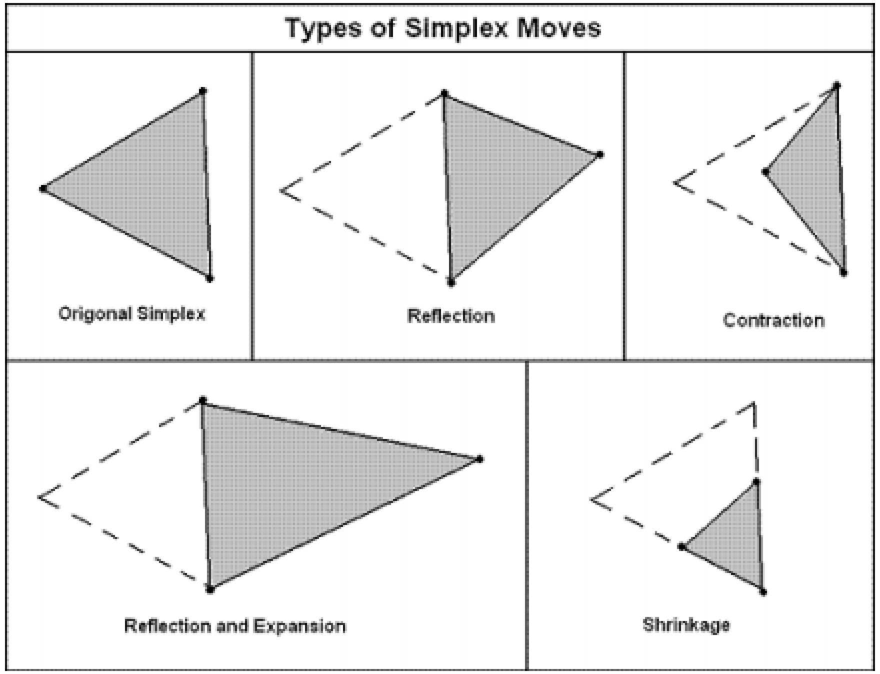
$$\chi^2 = \sum_{i=1}^{N} \left[\left(\frac{X_i - x_i}{\sigma_{X_i}} \right)^2 + \left(\frac{Y_i - y_i}{\sigma_{Y_i}} \right)^2 \right],$$

y에 대해서만 아닌 x에 대해서도 고려를 하여
y=ax+b 와 x=y/a-b/a
2개의 식을 이용하여 chi-square의 최소점을 구해본다

그렇다면 y에 대한 error만 있을 경우와 크게 차이가 없을거라고 생각이 되지만 a,b에 대한 error 값이 커져 그래프의 Chi-square fit가 넓어질것이다

기본원리

Downhill Simplex Method



Chi-Square 값이 가장 작은 부분을 찾아
Downhill Simplex 수행

실행

```python
if chre<=CH[0][0]:  #최저값이면 EXTENSION
    exx=2*rex-(CH[0][1]+CH[1][1])/2
    exy = 2 * rey - (CH[0][2] + CH[1][2]) / 2
    chex=ch(exx,exy)
    if chex<=chre:  #extension이 reflection 보다 작으면
        return CH[0][1], CH[1][1],exx, CH[0][2], CH[1][2], exy
    elif  chex>chre:  #extension이 reflection보다 크면
        return CH[0][1], CH[1][1],rex, CH[0][2], CH[1][2], rey
elif chre>CH[0][0] and chre<=CH[1][0]: #reflection 값이 중간값이면
    return CH[0][1], CH[1][1],rex, CH[0][2], CH[1][2], rey

elif  chre>CH[1][0]:  #reflection 한 값이 최고값이면 contraction
    cox=(CH[0][1]+CH[1][1]+CH[2][1])/2
    coy =(CH[0][2] + CH[1][2] + CH[2][2]) / 2
    chco=ch(cox,coy)
    if chco>CH[1][0]: #여전히 최고값이면 shrink 함
        return CH[0][1],(CH[1][1]+CH[0][1])/2, (CH[2][1]+CH[0][1])/2, CH[0][2],  (CH[1][2]+CH[0][2])/2, (CH[2][2]+CH[0][2])/2
    else: #아니면 contraction 반환
        return CH[0][1],CH[1][1], cox, CH[0][2], CH[1][2],  coy
```

Reflection 한 지점의 chi-square가 최저이면 extension 진행

Extension 한곳이 최저이면 그 좌표 반환 or 아니면 reflection 한 좌표 반환

Reflection 한 지점의 chi-squar가 최고이면 contraction 진행

Contraction 한 지점이 최고값이면 shrink 진행

```python
def downhill(x1,x2,x3,y1,y2,y3):
    p=[x1,x2,x3,y1,y2,y3]
    while 1:
        if np.abs(p[0] * p[4] + p[3] * p[2] + p[5] * p[1] - p[4] * p[2] - p[3] * p[1] - p[0] * p[5]) / 2 < 0.0001 :
            return p
            break
        else:
            p=determine(p[0],p[1],p[2],p[3],p[4],p[5])
```

세 점의 삼각형 넓이가 일정 이하까지 내려갈때가지 진행

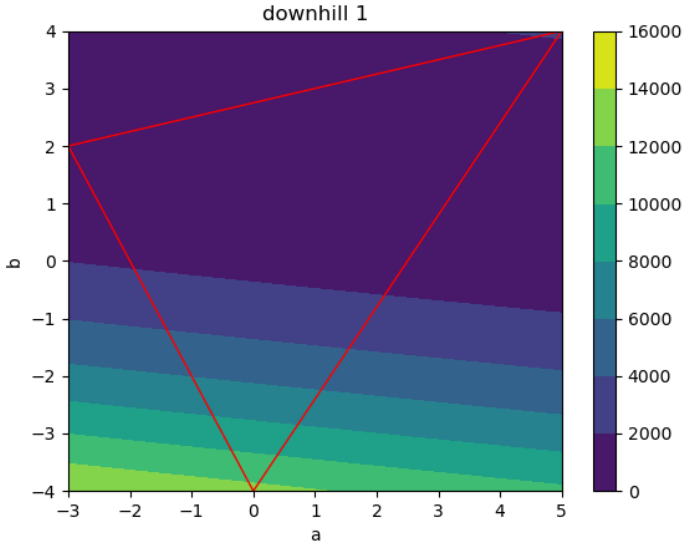
결과/토의사항

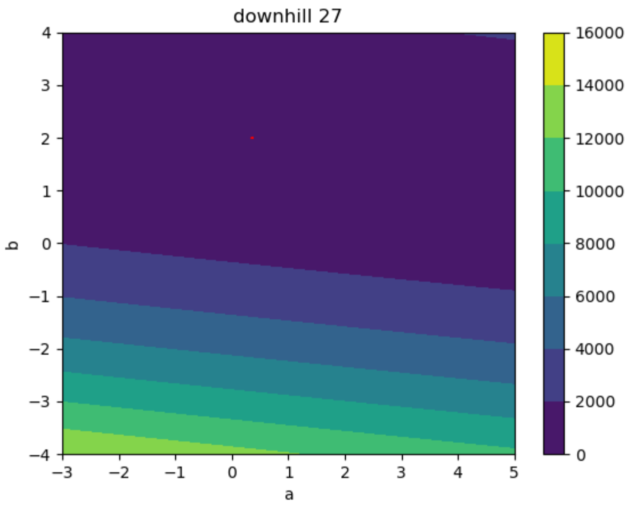
```
=downhill(5,0,-3,4,-4,2)
```

초기 삼각형의 좌표를
(5,4) , (0,-4), (-3,2) 로 설정후

Downhill 결과(삼각형 길이 수렴조건)

```
0.37712227925658226 1.9939513327553868
0.37712388671934605 1.9939508619718254
0.37712831888347864 1.9939506954979151
```





과제 8-1의 결과값과 비교

a값	0.37712438758161604
b값	1.9939510767614683
a오차	0.8438718500969482
b오차	0.10806292881858451

8-1 과 8-2 의 a,b값은 거의 유사하지만 Error에 대해서는 8-2(Downhill)의 방법이 더 적은값이 나왔다.

오차 계산

$$\frac{\partial^2 f}{\partial a^2} \sim \frac{1}{\sigma_a^2}$$

$$f''(x) \approx \frac{f(x+h) - 2f(x) + f(x-h)}{h^2}$$

h=0.000001으로 설정 시
적절한 2차미분값 도출

a오차	0.29745481000563945
b오차	0.03809016683024241

토의사항

수렴조건?

수렴조건의 경우 3가지 경우를 두고 downhill을 실행 해보았다.

1.삼각형의 넓이를 기준으로
2.삼각형의 변의 길이를 기준으로
3.세 지점에서의 reduced chi square를 기준으로

첫번째와 두번째의 경우 수렴기준의 값을 얼마나 작게 설정하냐에 따라 값이 달라질수 있지만 매우 작게 설정한다면 정상적으로 downhil의 결과가 나옴을 알수 있다.

Ex) (3,5) (1,2) (2,3) 에 대하여 삼각형의 넓이를 기준으로 downhill 진행
기준을 넓이 0.0001 이하로 진행시

```
0.96875 1.9375
0.9609375 1.921875
0.9765625 1.9453125
```

기준을 넓이 0.00000001 이하로 진행시

```
0.37750167958438396 1.9939096989110112
0.3775776815600693 1.9939235306810588
0.3768957192078233 1.9939954816363752
```

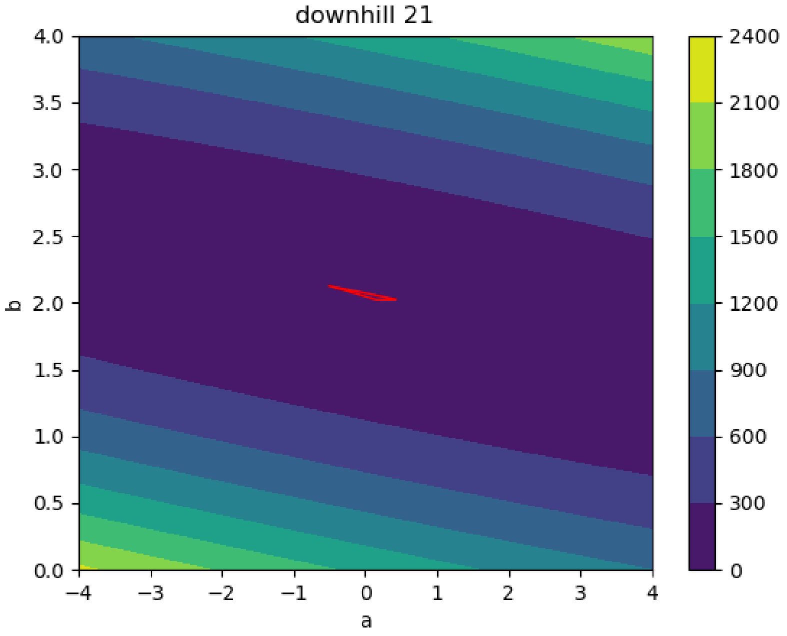
초기조건에 따라 매주 작은 수렴 조건 필요

토의사항

수렴조건?

세번째 경우는 각각 지점에 Chi square가
12 미만일때까지를 수렴조건으로 해보았다
초기좌표 (4,0) (-4,2) (-4,4)

```
0.4375 2.0234375
0.15625 2.021484375
-0.515625 2.1279296875
```



Reduced chi square 값?

```
0.6001758574846698
0.6001758574862646
0.6001758574860314
```

세 지점 모두 1이하로
적절하게 fitting 되었다

세 지점의 Chi square가 수렴조건 이하가 될 경우 downhill이 끝나 삼각형이 적절히 작아지지 않음을 볼수 있다

첫번째 두번째의 수렴조건을 사용하되 적절하게 작은 수렴기준이 필요함을 볼수있다.

토의사항

초기조건?

초기조건을 (3,4) (1,2) (2,3) 과 같이 삼각형이 아닌 초기 조건을 설정하였을때 다음과 같이 나옴을 볼수있다

```
0.9304924011230469 1.9304924011230469
0.9304904937744141 1.930490493774414
0.9304943084716797 1.9304943084716797
```

a,b를 `randrange(-100,100)` 로 설정후
downhill 실행결과 대부분 경우에서
원하는 값으로 downhill이 실행이 되었다.

삼각형을 이루는 좌표로 초기조건 설정

경계를 벗어나는 경우?

우선 결과값을 도출할때에는 경계를 벗어날 걱정은 없다.

하지만 plot 하여 gif 를 만들 경우 평면의 경계를 어떻게 설정할 것인가?

Downhill 실행 과정 도중 나온 삼각형 좌표중
최소, 최대값을 저장하여
이를 범위로 평면 경계를 설정한다

```
if max([p[0],p[1],p[2]])>maxx: maxx=max([p[0],p[1],p[2]])
if max([p[3],p[4],p[5]]) > may: may = max([p[3],p[4],p[5]])
if min([p[0],p[1],p[2]])<mix: mix=min([p[0],p[1],p[2]])
if min([p[3],p[4],p[5]]) <miy: miy = min([p[3],p[4],p[5]])
```

Reference

https://www.physicsforums.com/threads/chi-squared-fit-with-errors-on-both-x-and-y.980490/

Peter L. Jolivette, least-squares fits when there are errors in X