自然科学の歩き方

第4回目

前回やったこと

電流Iと電圧Vの間に$I=aV$という関係があるものと推定し，最適なaの値を探すことを考える。

V [V]	1.50	3.00	4.50	6.00	7.50	9.00
I [A]	5.64×10^{-2}	1.12×10^{-1}	1.86×10^{-1}	2.22×10^{-1}	3.25×10^{-1}	3.32×10^{-1}

1. aの値を0.02から0.01刻みで0.06まで変化させる。それぞれのaの値について，二乗誤差Eを計算し，表に書き込め。

aの値	0.02	0.03	0.04	0.05	0.06
Eの値					

2. 今回調べたaの値の中で，最も良く測定データを表すものは何か。そのaを用いた$I=aV$の直線を，グラフ中に実線で書き込め。

今回の話題

- ⭐ モデルパラメータ推定の定式化
  - ⭐ 「良いモデル」が何かをどう決めるか（簡単な紹介）
- ⭐ 「最小二乗法」の考え方と定式化

良いモデルとは（復習）

- ⭐ パラメータの数が少ない
  - ⭐ 予言能力が高い
- ⭐ 実験データをそれなりに再現できる
  - ⭐ 二乗誤差が小さい
- ⭐ 理論的根拠がしっかりしている

測定データの解釈

V [V]	1.50	3.00	4.50	6.00	7.50	9.00
I [A]	5.64×10^{-2}	1.12×10^{-1}	1.86×10^{-1}	2.22×10^{-1}	3.25×10^{-1}	3.32×10^{-1}

- 簡単なモデルとして，電圧Vと電流Iの間に比例関係が成り立つと仮定する $I = aV$
  - グラフからの直感
  - オームの法則
- aの最適な値は何か？

測定データの解釈

1. aの値を１つ決めて二乗誤差を計算する

- $$(\text{二乗誤差}) = \sum \left[(\text{モデルの予言値}) - (\text{データの値})\right]^2$$
- 二乗誤差はモデルとデータがどれくらい一致しているかを表す指標

2. aの値を変えながら，二乗誤差を計算する

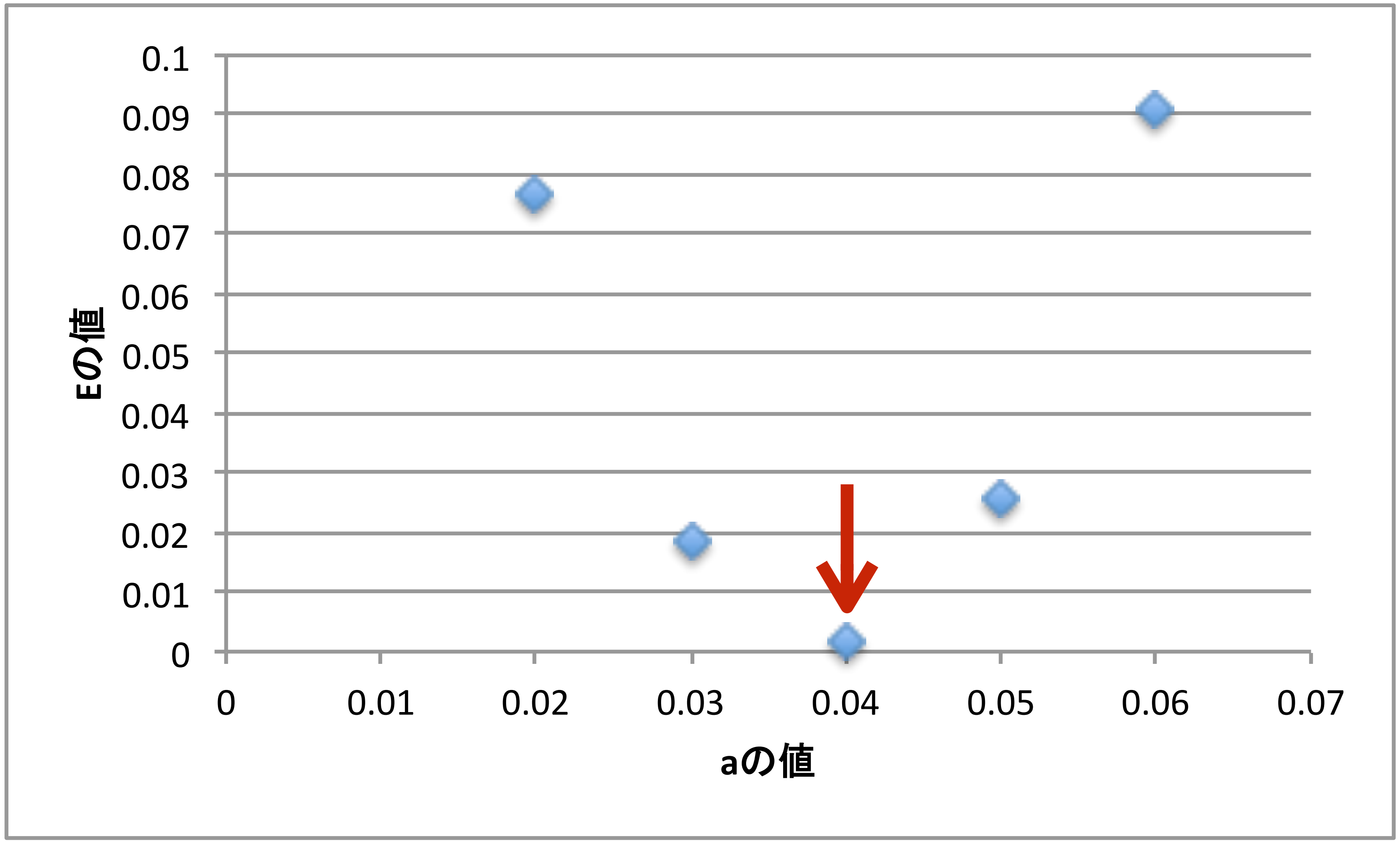
- 二乗誤差が小さいほどモデルとデータの一致度が高い

aの値	0.02	0.03	0.04	0.05	0.06
Eの値					

どれが一番小さいか？

結果

aの値	0.02	0.03	0.04	0.05	0.06
Eの値	0.0767	0.0188	0.00185	0.0258	0.0907



もっと効率よく解析する

- aを固定しないままで，二乗誤差を式で表してみよう
  - 二乗誤差をEとする
  - 電圧と電流の**測定値**の組を(V_i, I_i)で表す

V_1 V_2 $\cdots$ V_6

V [V]	1.50	3.00	4.50	6.00	7.50	9.00
I [A]	5.64×10^{-2}	1.12×10^{-1}	1.86×10^{-1}	2.22×10^{-1}	3.25×10^{-1}	3.32×10^{-1}

I_1 I_2 $\cdots$ I_6

$$E = \sum_{i=1}^{6} \left(I_{\text{モデル}} \Big|_{V=V_i} - I_i \right)^2 = \sum_{i=1}^{6} (aV_i - I_i)^2$$

この式の読み方

- ⭐ 測定値の値は決まっている（変えてはいけない）
- ⭐ 変えていいのは，$I=aV$の直線の傾きのa
- ⭐ つまりaは変数とみなせる（Eはaの関数）

$$E(a)=\sum_{i=1}^{6}\left(aV_i-I_i\right)^2=\left(aV_1-I_1\right)^2+\left(aV_2-I_2\right)^2+\cdots\left(aV_6-I_6\right)^2$$

$$=a^2\left(V_1^2+\cdots+V_6^2\right)-2a\left(V_1I_1+\cdots+V_6I_6\right)+\left(I_1^2+\cdots+I_6^2\right)$$

二乗誤差を最小化する

- Eはaの関数
- Eを最小化する→$E(a)$の最小値を求める
  - $E(a)$をaで微分して0になる点（極小値）を探す $\dfrac{dE}{da}=0$
    aの定義域の端が最小の可能性もある
  - (今のような二次関数の場合は，平方完成して頂点を求めてもよい。)

$$\frac{dE}{da} = 2a\left(V_1^2 + \cdots + V_6^2\right) - 2\left(V_1 I_1 + \cdots + V_6 I_6\right) = 0$$

微分とは何か（復習）

- ☆ 変数xについての関数y=$f(x)$を考える。xが微小に変化した時，関数$f(x)$の値がどのように変化するか？

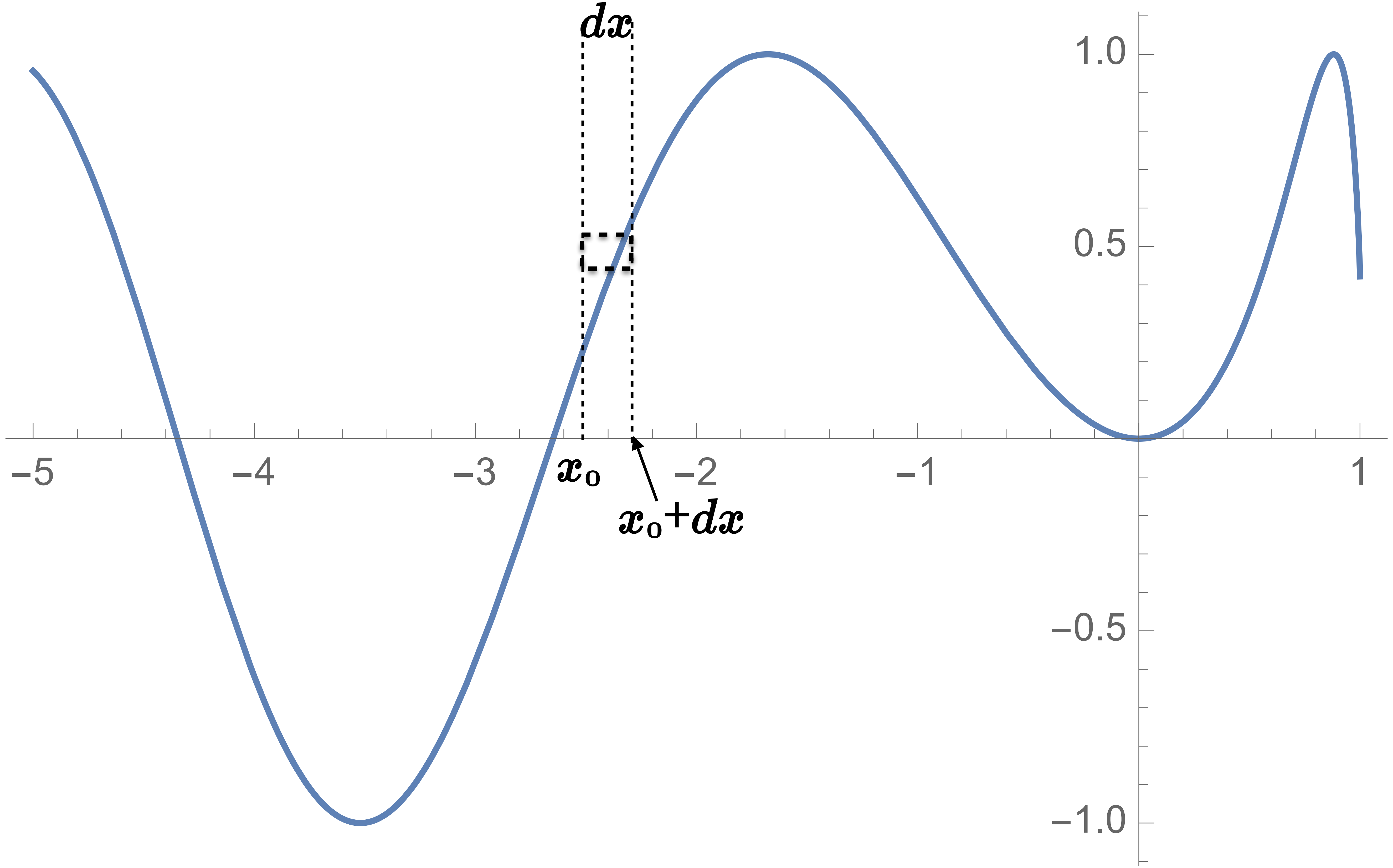
- ☆ 正しくはxやyの「微小変化分」dx, dyのことを微分という。

- ☆ xがΔxだけ変化した時，

$$\Delta f = f(x + \Delta x) - f(x)$$

Δxを無限に小さくしていくと，どうなるか？

十分小さくとったΔxをdxと書く。

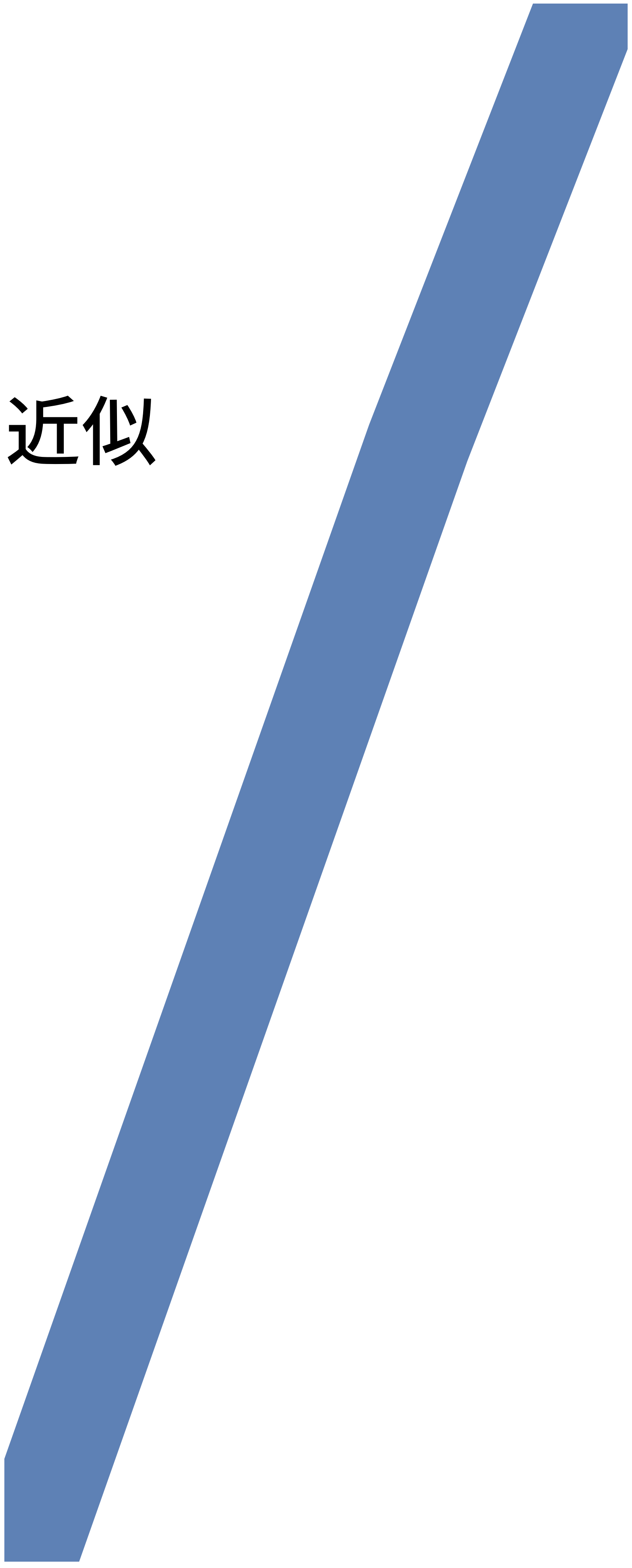


dx
1.0
0.5
−0.5
−1.0
−5
−4
−3
x_0
−2
−1
1
x_0+dx

直線で近似できる！

$$y = ax + b$$

ただし，$\boldsymbol{x_0}$~$\boldsymbol{x_0}$+$\boldsymbol{dx}$の区間のみで成り立つ近似

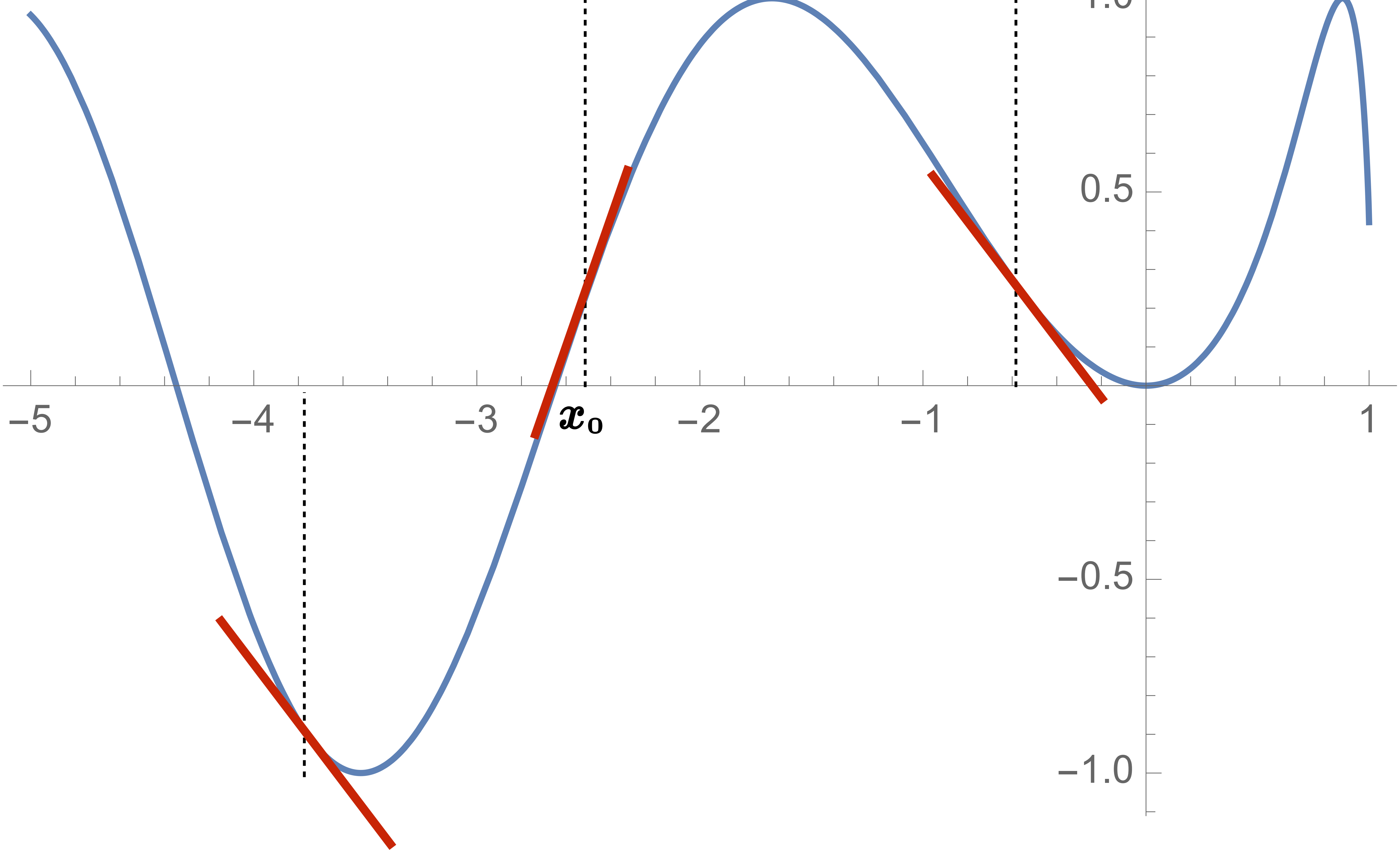


この直線の傾きは？

$$\frac{df(x_0)}{dx} = \lim_{\Delta x \to 0} \frac{f(x_0 + \Delta x) - f(x_0)}{\Delta x}$$

$\boldsymbol{x}$=$\boldsymbol{x_0}$における微分係数という

微分係数は接線の傾き



x_0を色々変えると，それに応じて微分係数の値も変化

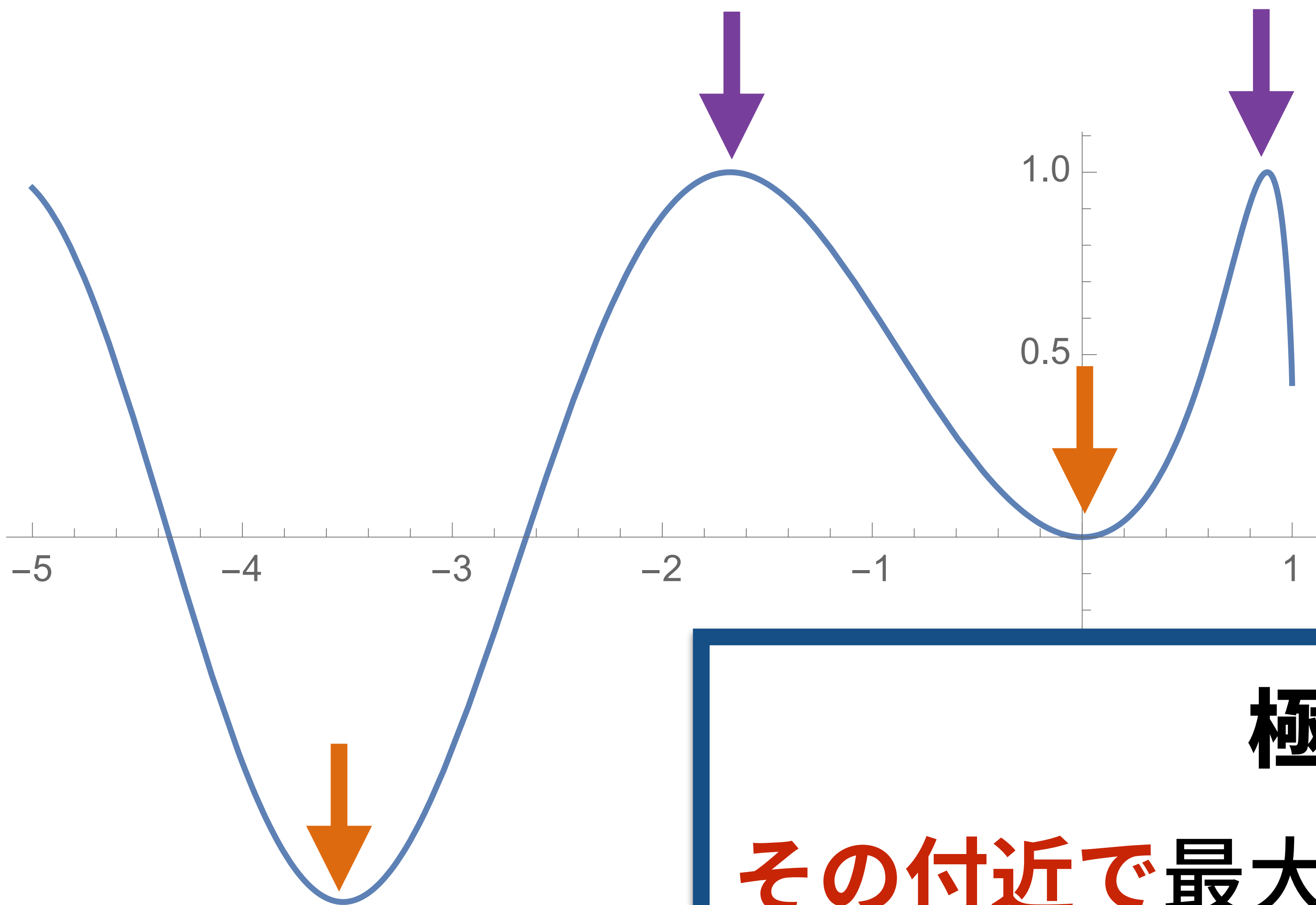
$$\frac{df(x)}{dx}$$:導関数

微分の利用その1

微分係数＝グラフの（接線の）傾き

↓

関数の極大，極小を求めるのに微分係数を利用できる

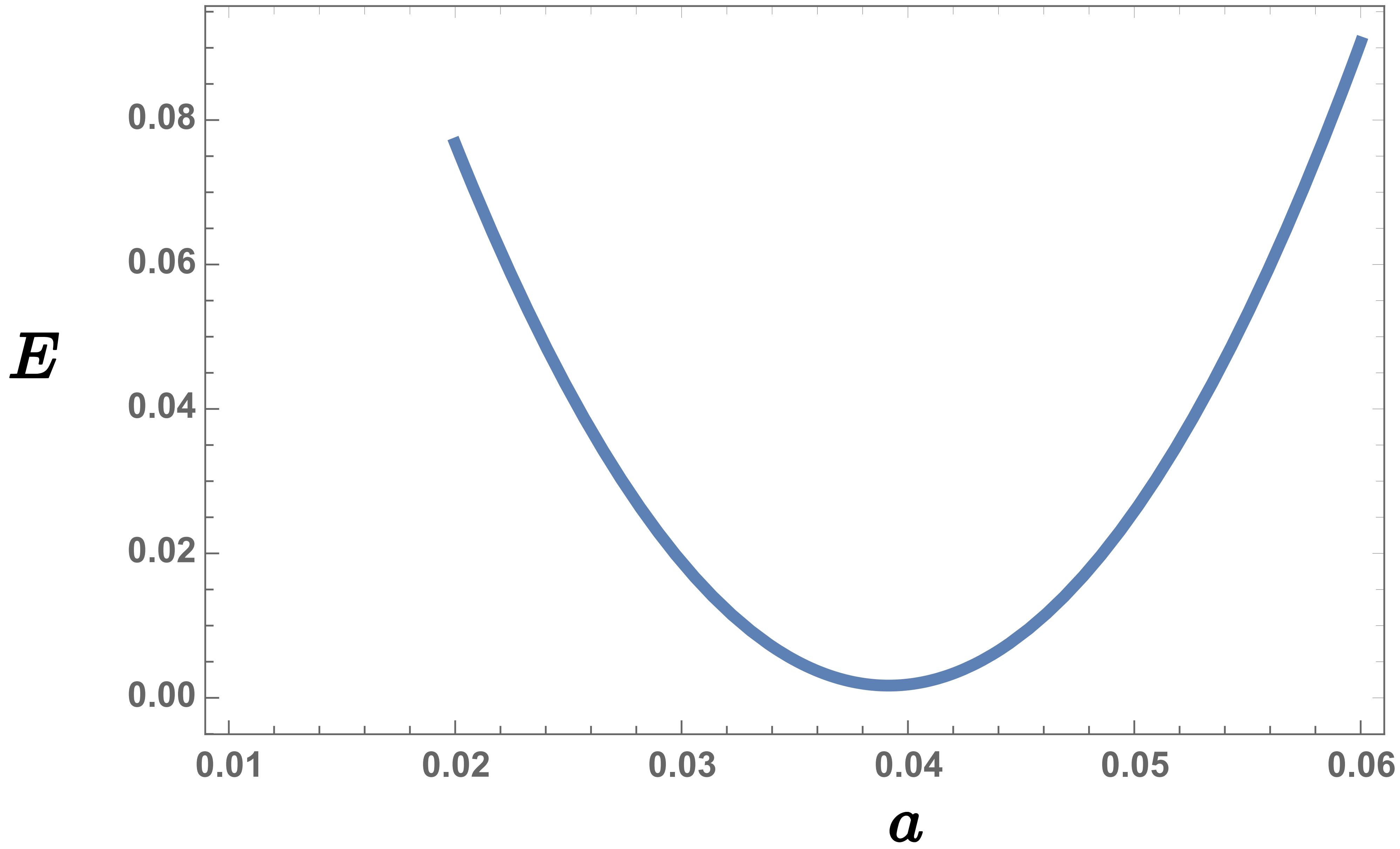


極大・極小

<u>その付近で</u>最大・最小の点

定義域内での最大・最小とは限らない

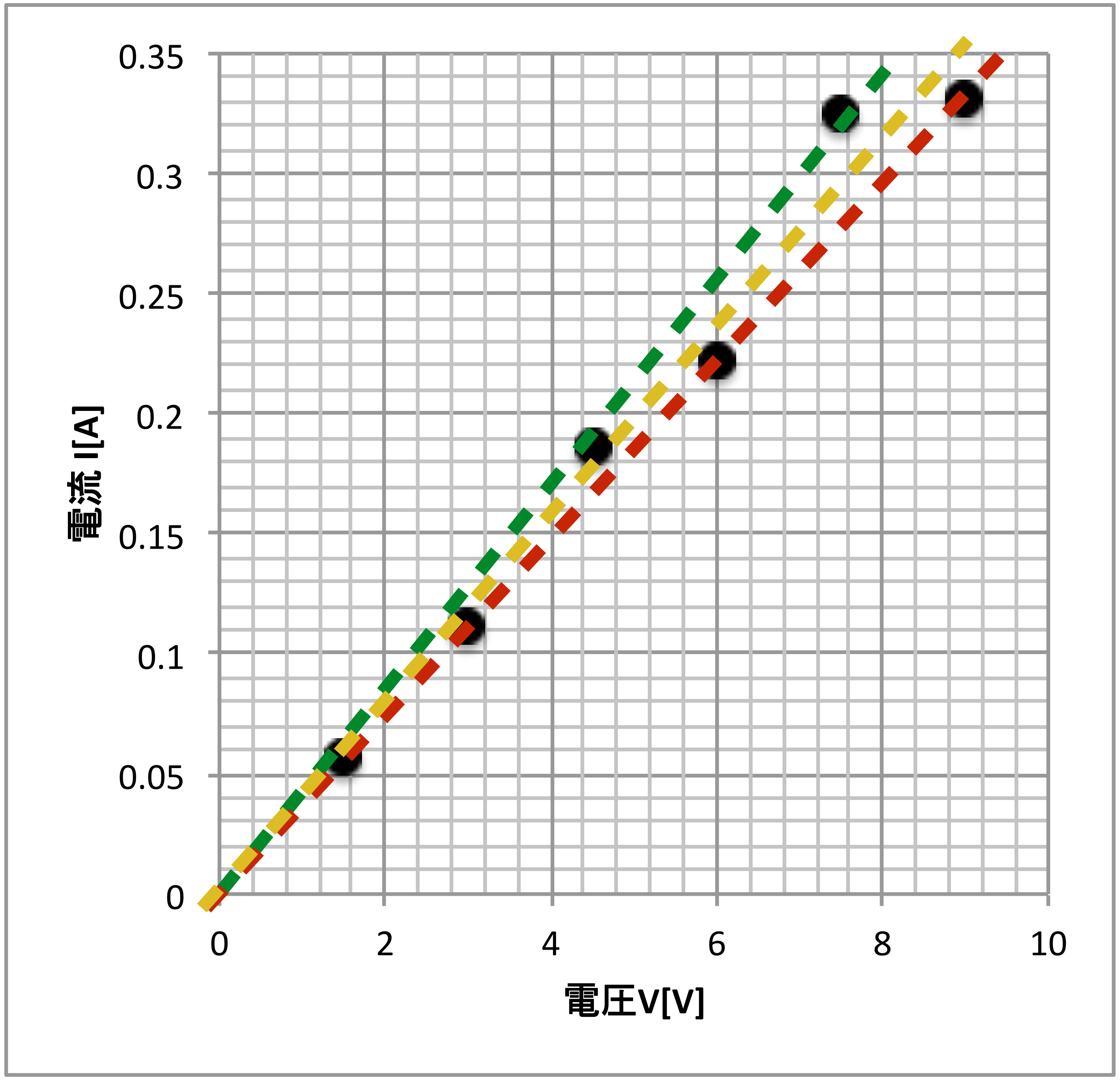
$E(a)$のふるまい



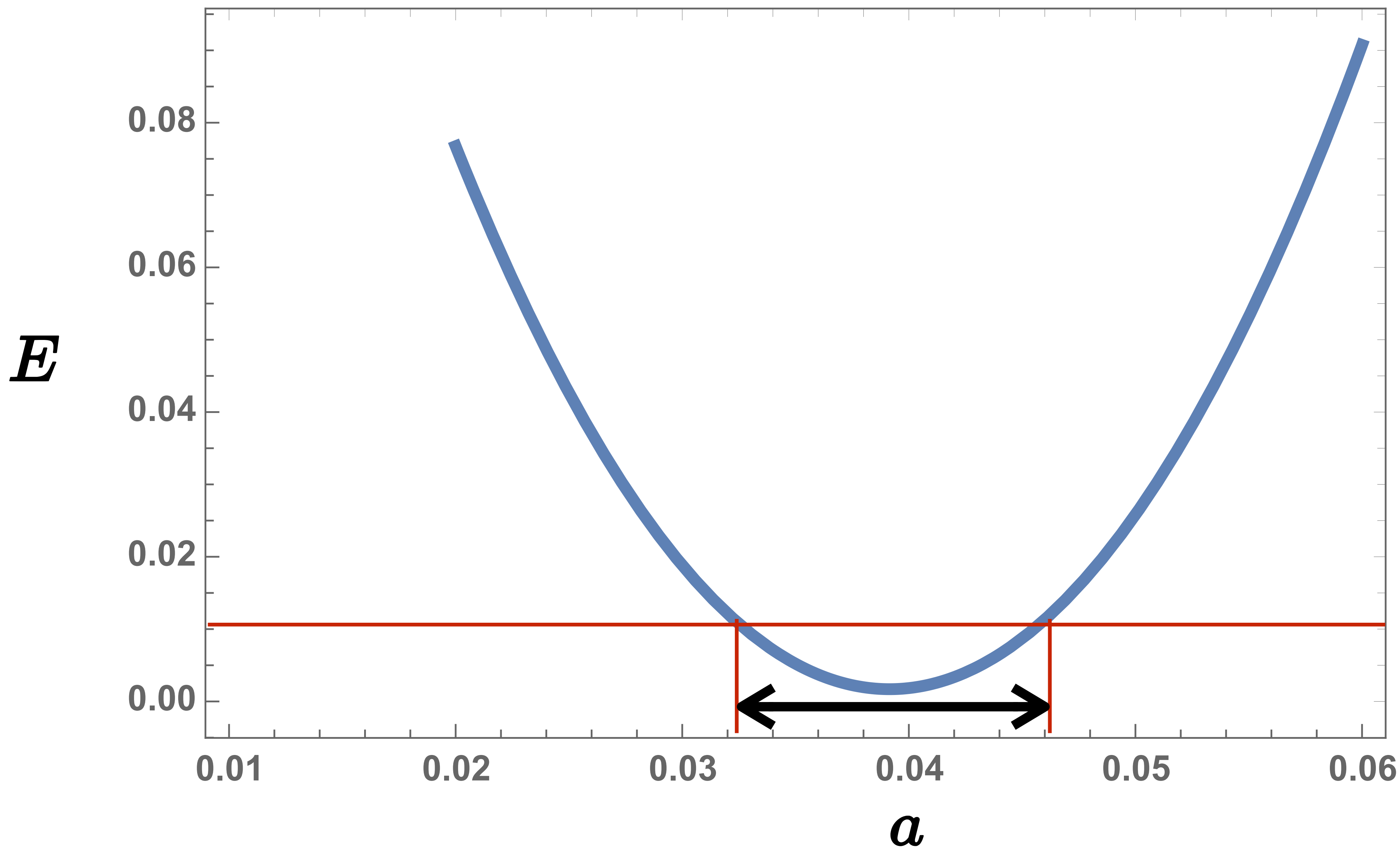
最適なaはEが最も小さいところ

パラメータ推定の誤差

- どうせモデルの線は測定点すべてを通らない
- 測定にも誤差がある
- データが追加されると最適なaの値も動く可能性がある
- パラメータの値も「厳密」に決まったりはしない



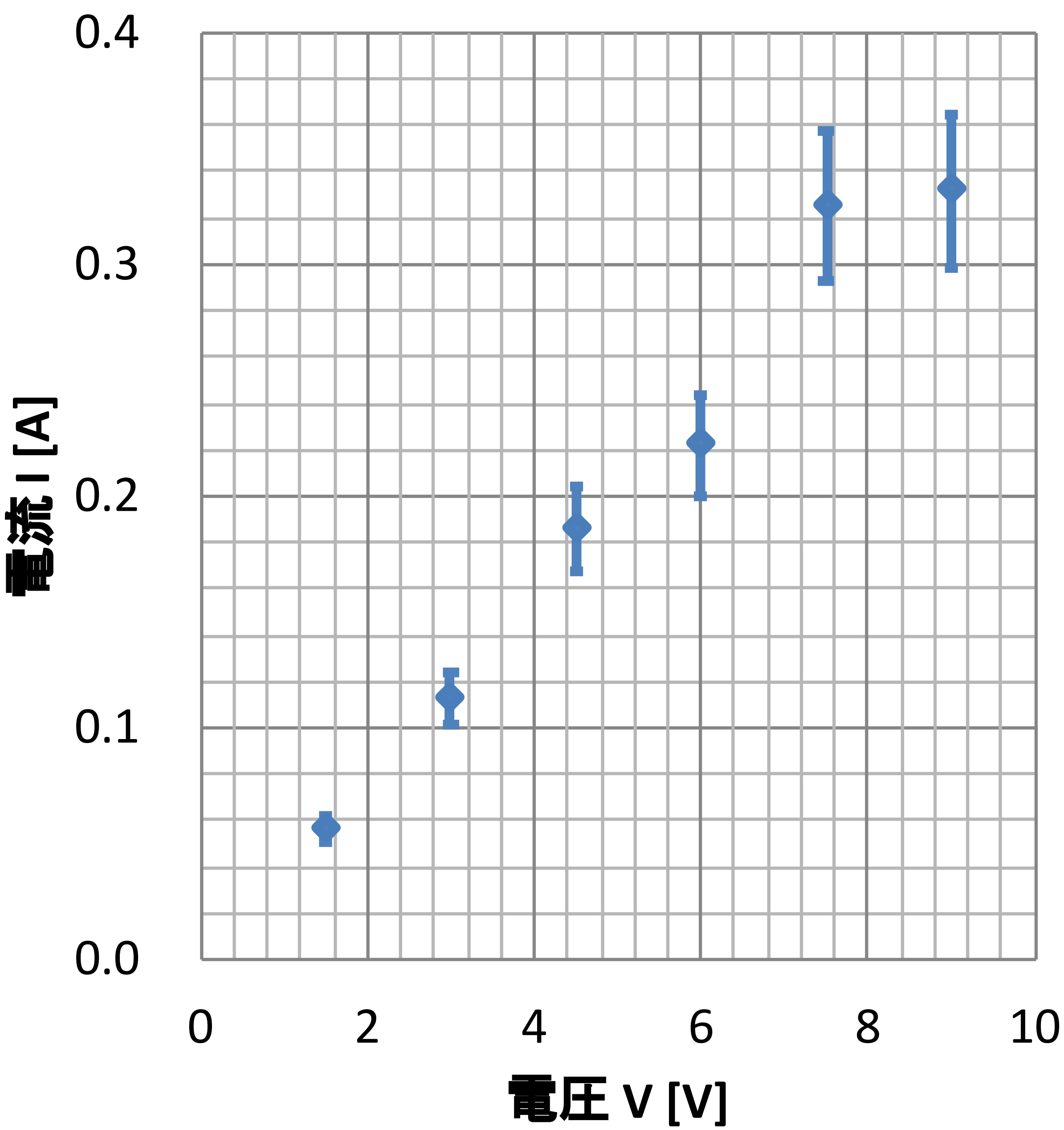
パラメータの誤差



二乗誤差がある値以下になるべしという条件を設定して，
パラメータの真の値が存在しそうな領域を決める

データには誤差がある

誤差の評価されていない測定データには意味がない



測定データは本来

$$I_i \pm \sigma_i$$

誤差

のように表されるべき

正式な最小二乗法

あるモデル$y=f(x)$が与えられた時，

測定データの組$(x_1,y_1),(x_2,y_2)\cdots$を用いて，

$$\chi^2 = \sum_i \frac{(y_i - f(x_i))^2}{\boxed{\sigma_i^2}}$$

測定や計算に伴う誤差²

が最小になるようにモデルパラメータを決める

簡単な場合として$\sigma_1 = \sigma_2 = \cdots = \sigma_N$ を考えると

$$\sigma^2\chi^2 = \sum_i (y_i - f(x_i))^2$$

二乗誤差の和

が最小のときが，最適なパラメータの値

χを最小にする意味

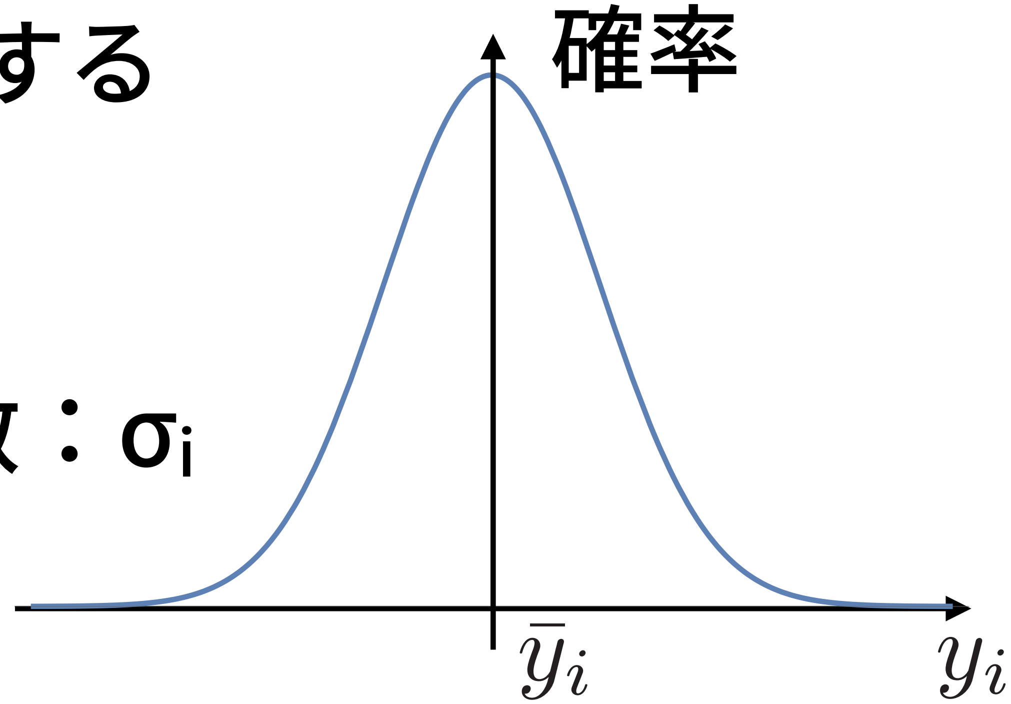
- ☆ 測定データの組$(x_1,y_1),(x_2,y_2)\cdots$があるとする

- ☆ どの測定も互いに影響しあわないとする（独立）

  - ☆ 同じx_iに対して何度も実験をした場合，測定値yᵢのばらつき具合は正規分布に従うとする

$$P(y_i;\bar{y}_i,\sigma_i)=\frac{1}{\sqrt{2\pi\sigma_i^2}}e^{-\frac{(y_i-\bar{y}_i)^2}{2\sigma_i^2}}$$

分散：σ_i



分散の意味：測定値が68.3%の確率で $\bar{y}_i\pm\sigma_i$ の範囲内に現れる

95% $\bar{y}_i\pm 2\sigma_i$

99% $\bar{y}_i\pm 3\sigma_i$

測定値の組を得る確率

モデル$y=f(x)$が正しいとする→y_iの測定平均は$f(x_i)$になるはず

1個の観測値(x_i,y_i)を得る確率は

$$P(y_i;\bar{y}_i,\sigma_i)=\frac{1}{\sqrt{2\pi\sigma_i^2}}e^{-\frac{(y_i-\bar{y}_i)^2}{2\sigma_i^2}}$$

N個の観測値の組{(xᵢ,yᵢ),i=1,...,N}が得られる確率は

$$P[(x_1,y_1),(x_2,y_2),\cdots,(x_N,y_N)]=P[(x_1,y_1)]P[(x_2,y_2)]\cdots P[(x_N,y_N)]$$

$$=\frac{1}{\sqrt{(2\pi)^N\prod\sigma_i^2}}e^{\left[-\frac{1}{2}\boxed{\sum\frac{(y_i-f(x_i))^2}{\sigma_i^2}}\right]}$$

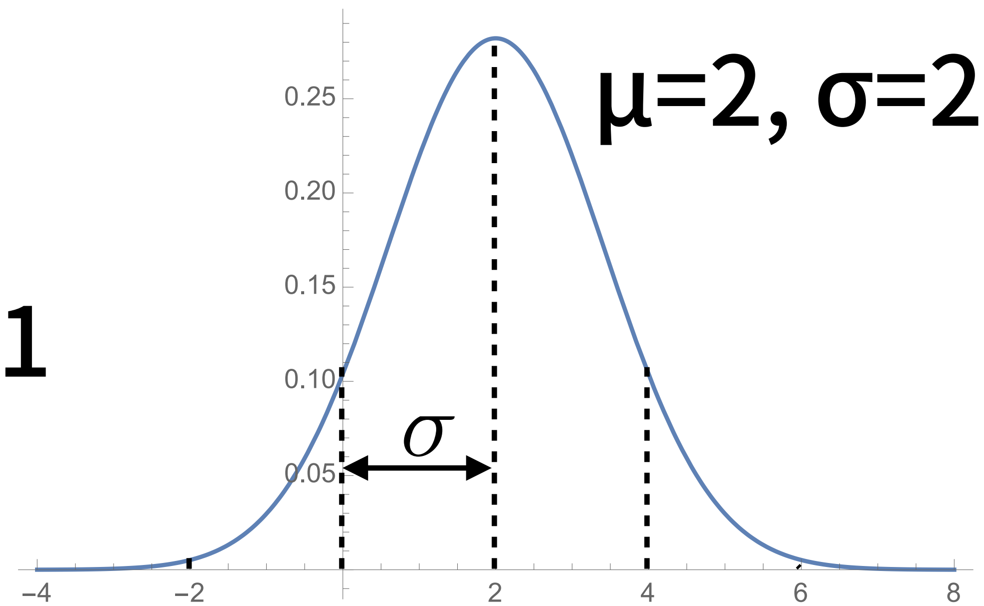
仮定したモデルにおいて，実際測定されるような観測値の組が与えられる確率が最大になるのは

の部分が最小になるときである

おまけとして検定の話

$$f(x) = \frac{1}{\sqrt{2\pi\sigma}} e^{-\frac{(x-\mu)^2}{2\sigma}}$$

$$\int_{-\infty}^{\infty} f(x)dx = 1$$ 確率の合計は1



a<x<bとなる確率は

$$P[a < x < b] = \int_a^b f(x)dx$$

- ⭐ -σ<x-μ<σとなる確率:68.3%
- ⭐ -1.64σ<x-μ<1.64σとなる確率:90%
- ⭐ -1.96σ<x-μ<1.96σとなる確率:95%
- ⭐ -2σ<x-μ<2σとなる確率:95.5%
- ⭐ -2.58σ<x-μ<2.58σとなる確率:99%
- ⭐ -3σ<x-μ<3σとなる確率:99.7%

よくやる使い方

- ☆ 物理量xを測定し，平均値が2,　統計誤差が1となった
- ☆ 仮説H: xの真の値は0である。
  - ☆ 対立仮説H’: xの値は0ではない（本当はこれを示したいことが多い）
  - ☆ 実はH’を示したいような場合，Hを無帰仮説とよぶ
- ☆ Hが定められた危険率αより大きいか小さいかで検定
  - ☆ 今の例だと，Hは2σ(95.5%)で棄却できる（危険率4.5%)

全ての不可能を消去して，最後に残ったものがいかに
奇妙であったとしても，それが真実である

by シャーロックホームズ

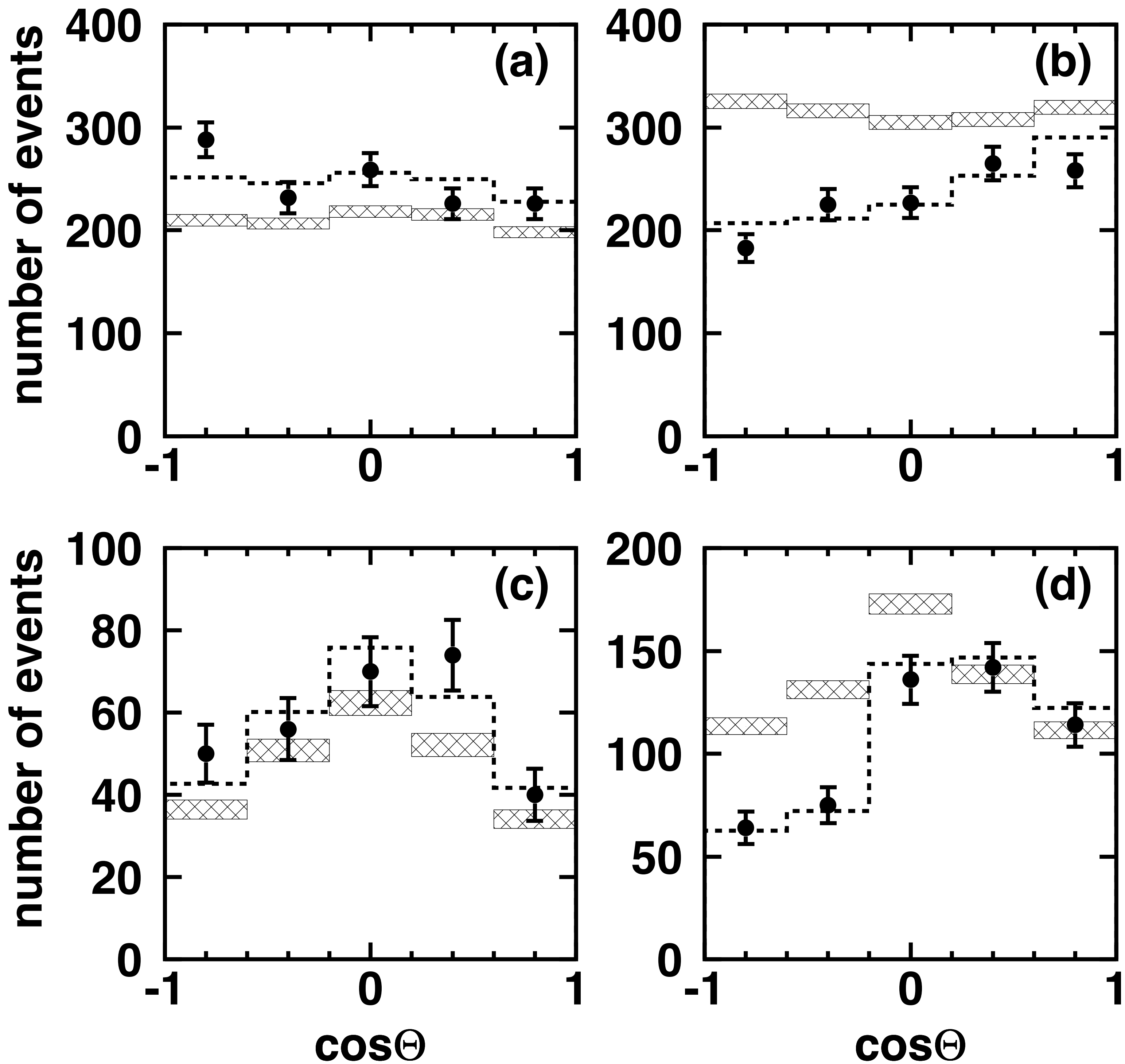


Figure 1. Zenith angle distributions observed in Super-Kamiokande for; (a)sub-GeV e-like, (b)sub-GeV μ-like, (c)multi-GeV e-like and (d) multi-GeV (FC+PC) μ-like events. CosΘ =1 means down-going particles. The histograms with shaded error bars show the MC prediction with their statistical errors for the no neutrino oscillation case. The dotted histograms shows the Monte Carlo prediction for $\nu_\mu \leftrightarrow \nu_\tau$ oscillations with $\sin^2 2\theta$ =1 and Δm^2 =2.2×10^{-3} eV^2.

hep-ex/9810001より

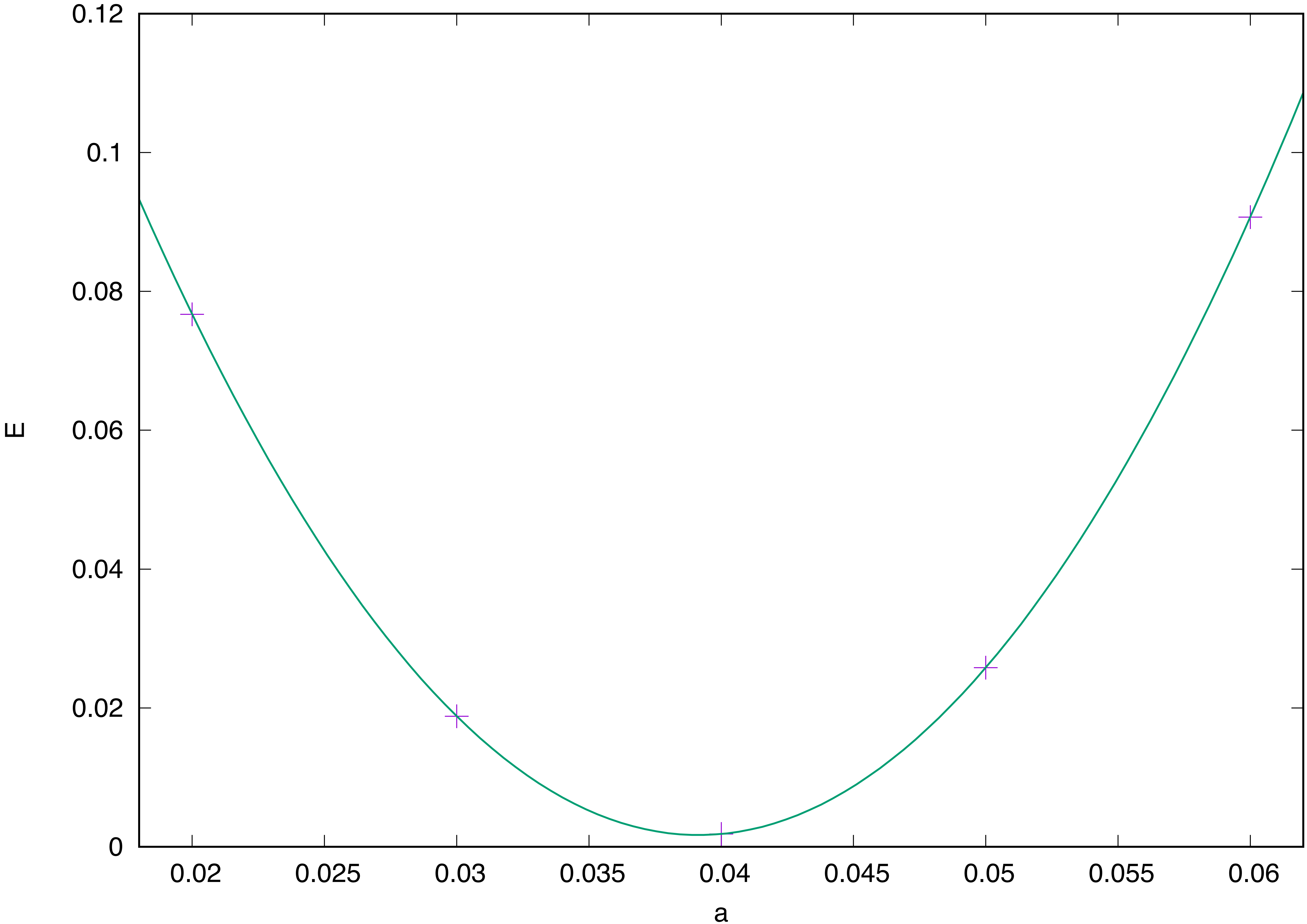
今日の演習問題

1. 前々回に作成した，aとEの表をグラフに示せ。
（点が5個描かれたグラフができる）

2. Eを，定義式をもとにして，aの関数として式で表し，そのグラフの概形を，上で作成したグラフ上に実線で示せ。

3. Eを最小にするaの値を，式に基づいて計算せよ。
（たいていは，２の段階ですでに求まっている）

$$E = \sum_{i=1}^{6} (aV_i - I_i)^2$$


E
a
0
0.02
0.04
0.06
0.08
0.1
0.12
0.02
0.025
0.03
0.035
0.04
0.045
0.05
0.055
0.06