

『AI・データサイエンスのための数学入門』学習補助資料

変化の形から、モデルを選ぶ

一次・二次・指数・周期モデルとデータへの当てはめ

1 「一定量」と「一定割合」を区別する

モデルを選ぶときには、グラフの形に加えて、入力と同じだけ進んだとき、何が一定になるかを確認します。

候補	基本形	注目する変化
一次	$ax + b$	同じ入力増分に対する出力の差
二次	$ax^2 + bx + c$	等間隔のデータにおける差の差
指数	Ca^x	同じ入力増分に対する出力の比
周期	$A \sin(\omega x + \phi) + c$	繰り返しの間隔と振幅

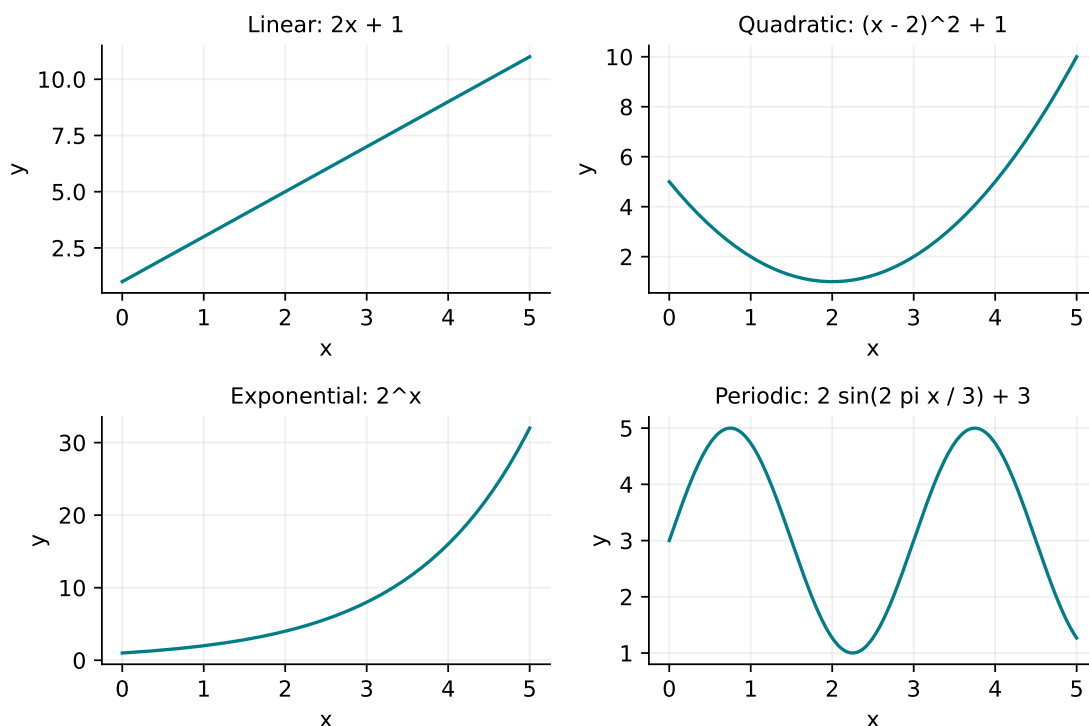


図 1：4 種類のモデルの例。各図は縦軸の範囲が異なるため、曲線の見た目の傾きだけで増加の速さを比較しないようにします。

モデルは現象を、ある範囲で近似するもの

観測値は一般に $y_i = f(x_i) + \varepsilon_i$ と考えます。基本モデルに完全一致しなくても、傾向を記述する助けになります。ただし、観測範囲の外でも同じ法則が続くとは限りません。

例えば普及の初期には指数的に増えていても、対象となる人数に上限があれば、その増加は永続できません。モデルを使う期間と入力範囲を明確にします。

2 一次関数：差と単位から意味を読む

$f(x) = ax + b$ に対して、任意の増分 h について

$$f(x+h) - f(x) = a(x+h) + b - (ax+b) = \boxed{ah}$$

です。 a は「入力 1 単位あたりの出力の変化量」を表します。

例：固定費と変動費

総費用が $C(q) = 500 + 120q$ 円なら、500 円が固定費、120 円/個が単価です。生産を 3 個増やすと、元の生産量にかかわらず費用は 360 円増えます。

一方、平均費用は $C(q)/q = 500/q + 120$ ($q > 0$) であり、一次関数ではありません。総量と 1 個あたりの量を区別します。

比例とアフィンの違い

ax は原点を通り、 $f(x_1 + x_2) = f(x_1) + f(x_2)$ を満たします。 $ax + b$ では、

$$f(x_1) + f(x_2) = a(x_1 + x_2) + 2b$$

となり、 $b \neq 0$ なら加法性を満たしません。なお $a = 0$ の場合は定数関数であり、厳密には一次関数ではありません。

2 点から直線を決める

異なる入力 x_1, x_2 に対応する値が y_1, y_2 なら、

$$a = \frac{y_2 - y_1}{x_2 - x_1}, \quad b = y_1 - ax_1.$$

例えば $(2, 7), (5, 16)$ を通る直線は $a = 3, b = 1$ より $y = 3x + 1$ です。ただし、ノイズを含む多数の点に当てはめる場合には、すべての点を通る直線は通常ありません。

切片を現実の初期値と読めるか

切片は数式上の $x = 0$ の値です。しかし、観測した x が 50~100 だけなら、 $x = 0$ の値は遠い外挿です。切片に現実的な意味があるかは、入力の定義と適用範囲によります。

期待値と一次関数

X の期待値が存在すれば、期待値の線形性から

$$E[aX + b] = aE[X] + b.$$

料金のような一次式では、平均使用量を代入すれば期待料金になります。非線形な関数では、一般に $E[f(X)] \neq f(E[X])$ です。

3 二次関数：変化する増分と頂点

$f(x) = ax^2 + bx + c$ ($a \neq 0$) では、

$$f(x+h) - f(x) = 2axh + ah^2 + bh.$$

増分は入力 x にも依存します。さらに、同じ幅 h で取った差の差は

$$f(x+2h) - 2f(x+h) + f(x) = 2ah^2$$

となり一定です。これが、等間隔のデータで二次モデルを考える手がかりです。

差分の例： $y = x^2$

x	0	1	2	3	4
x^2	0	1	4	9	16
隣との増分		1	3	5	7
増分の差			2	2	2

平方完成で最適値を読む

$$ax^2 + bx + c = a \left(x + \frac{b}{2a} \right)^2 + c - \frac{b^2}{4a}.$$

二乗は非負なので、全実数上では $x = -b/(2a)$ で、 $a > 0$ なら最小、 $a < 0$ なら最大になります。微分を使わずに頂点と最適値が分かります。

頂点が許される範囲にあるか確認する

例えば $f(x) = (x-3)^2 + 1$ の全実数上の最小点は $x = 3$ ですが、 $0 \leq x \leq 2$ という制限があれば最小点は $x = 2$ 、最小値は 2 です。頂点を計算しただけで終わらず、定義域も確認します。

個数のように x が整数なら、実数上の最適点の近くにある許される整数を比較します。例えば頂点が 2.4 なら、制約も確認したうえで 2 と 3 の値を比較します。

「入力に非線形」と「係数に線形」

$ax^2 + bx + c$ は入力 x に対して二次式ですが、係数 a, b, c には一次式です。このため、多項式を当てはめる方法も、係数についての線形回帰として扱えます。何に関して線形なのかを区別します。

4 指数関数：一定割合の変化を積み重ねる

初期値 $C > 0$ を含めると、モデルは $y_t = Ca^t$ です。 $a > 0$ で

$$\frac{y_{t+1}}{y_t} = a, \quad \frac{y_{t+1} - y_t}{y_t} = a - 1.$$

$a = 1.05$ は 1 期間あたり 5% 増、 $a = 0.9$ は 10% 減です。減少する場合も値は正のままで、有限時刻に 0 になるわけではありません。

一定量と一定割合の比較

t	0	1	2	3
每期 20 増： $100 + 20t$	100	120	140	160
每期 20% 増： $100(1.2)^t$	100	120	144	172.8

倍加時間と半減期

$Ca^T = 2C$ を解くと、 $a > 1$ の倍加時間は

$$T = \frac{\log 2}{\log a}.$$

$0 < a < 1$ の半減期は $T = \log(1/2)/\log a$ です。これは連続時間に拡張したモデルでの時間であり、整数時点だけ観測するなら、目標を初めて超える時点などに読み替えます。

自然指数の形との対応

$a^t = e^{t \log a}$ なので、 $Ca^t = Ce^{rt}$ と書くなら $r = \log a$ です。離散的な増加率 $a - 1$ と r は同じではありません。例えば $a = 1.1$ では、増加率は 0.1 ですが $r = \log 1.1 \approx 0.09531$ です。

負の指数と分数指数を確認する

$$2^{-3} = \frac{1}{8}, \quad 16^{1/2} = 4, \quad 27^{2/3} = 9.$$

$a > 0$ という条件の下で、指数法則が実数の指数にも整合するように定義されています。

指数減衰するのは、どの量か

每期 10% が離脱するモデルでは、指数的に減るのは残っている人数 $N_t = N_0(0.9)^t$ です。この例の離脱率 10% は一定です。「率」と「人数」を混同しないようにします。

5 三角関数：周期・振幅・位相を分ける

ラジアンは、円弧の長さを半径で割った角度です。1 周が 2π 、半周が π 、直角が $\pi/2$ です。単位円では座標が $(\cos \theta, \sin \theta)$ となります。

$$\theta = \frac{\pi}{3} : \quad \cos \theta = \frac{1}{2}, \quad \sin \theta = \frac{\sqrt{3}}{2}, \quad \tan \theta = \sqrt{3}.$$

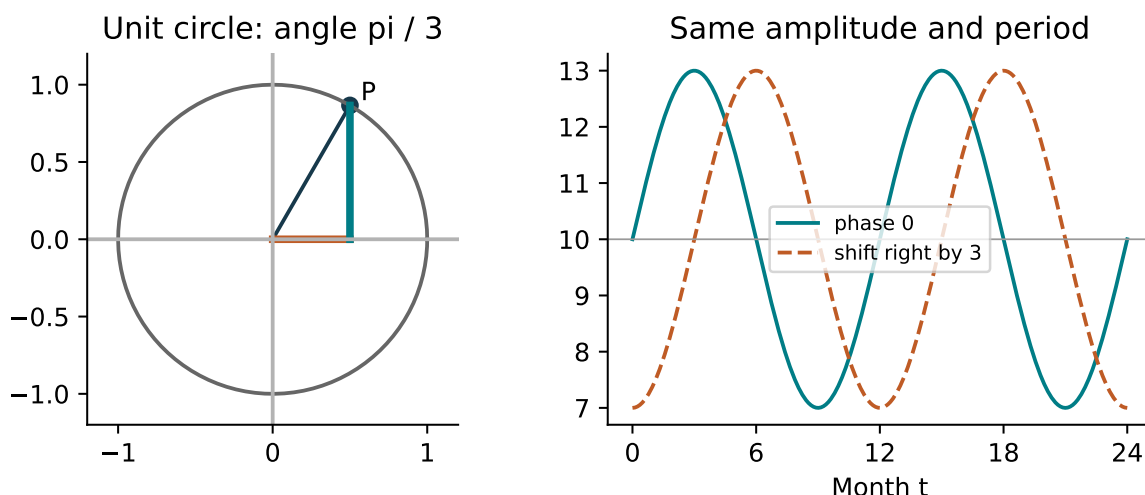


図 2：左は角度 $\pi/3$ の単位円。横の成分が余弦、縦の成分が正弦です。右は同じ振幅 3・周期 12・中心 10 を持ち、位相だけが違う波です。

周期モデルのパラメータを読む

$y = A \sin(\omega t + \phi) + c$ で $A \neq 0, \omega > 0$ とします。

性質	値
中心線	$y = c$
振幅	$ A $ (通常は $A > 0$ として指定)
最大値・最小値	$c + A , c - A $
周期	$T = 2\pi/\omega$
横方向の移動	$-\phi/\omega$ (周期を除いて考える)

$A > 0$ なら、山は $\omega t + \phi = \pi/2 + 2\pi\ell$ (ℓ は整数) の時点にあります。位相 ϕ は角度であり、時間のずれそのものではありません。

なぜ周期が $2\pi/\omega$ なのか

t が T 増えたとき角度が 2π 増えれば同じ値に戻るので、 $\omega T = 2\pi$ から求められます。 $A = 0$ の定数関数では、基本周期をこの式で決めるわけではありません。 $\tan \theta$ は $\theta = \pi/2 + \pi\ell$ で未定義です。正弦・余弦と同じように、全角度で有限な値を取るわけではありません。

6 本文・問 1、問 2 を解く

問 1：料金プランの比較

料金は

$$A(x) = 1800 + 450x, \quad B(x) = 3200 + 250x \quad (x \geq 0).$$

$B(x) < A(x)$ を解くと $1400 < 200x$ より $x > 7$ 。7GB なら同額、7GB 未満なら A が安くなります。

使用量	確率	A の料金	B の料金
2GB	0.2	2700 円	3700 円
6GB	0.5	4500 円	4700 円
10GB	0.3	6300 円	5700 円

平均使用量は $E[X] = 2(0.2) + 6(0.5) + 10(0.3) = 6.4$ なので、

$$E[A(X)] = 1800 + 450(6.4) = 4680 \text{円},$$

$$E[B(X)] = 3200 + 250(6.4) = 4800 \text{円}.$$

この分布の下で 1 つのプランを選ぶなら、期待料金では A が月 120 円有利です。毎月必ず A が安いという意味ではありません。

問 2：利益の最大化

売上は $p(q)q = (120 - 2q)q$ です。費用を引くと、

$$\begin{aligned} \Pi(q) &= (120 - 2q)q - (200 + 20q) \\ &= -2q^2 + 100q - 200 \\ &= -2(q - 25)^2 + 1050. \end{aligned}$$

$q \geq 0$ の範囲で $q = 25$ のとき最大となり、

$$p(25) = 120 - 2(25) = 70, \quad \Pi(25) = 1050.$$

経済的な範囲も確認する

価格が非負という条件を加えるなら $q \leq 60$ です。最適点 25 は $0 \leq q \leq 60$ に入るため、結論は変わりません。数量が整数でも 25 は許されます。

7 本文・問 3：当てはまりと予測を比較する

観測値は 100, 128, 166, 218、候補は $L(t) = 98 + 39t$ と $E(t) = 100(1.29)^t$ です。途中で丸めず計算すると、

t	観測	$L(t)$	観測 $-L(t)$	$E(t)$	観測 $-E(t)$
0	100	98	2	100	0
1	128	137	-9	129	-1
2	166	176	-10	166.41	-0.41
3	218	215	3	214.67	3.33

$$RSS_L = 2^2 + (-9)^2 + (-10)^2 + 3^2 = \boxed{194}.$$

指数モデルでは、正確には $E(2) = 166.41$, $E(3) = 214.6689$ なので、

$$RSS_E = 0^2 + (-1)^2 + (-0.41)^2 + (3.3311)^2 \approx \boxed{12.2643}.$$

問題文の近似値 166.4、214.7 を使えば 12.05 となります。丸め方で多少変わりますが、与えられた 2 候補では指数モデルの RSS が小さいという判断は同じです。

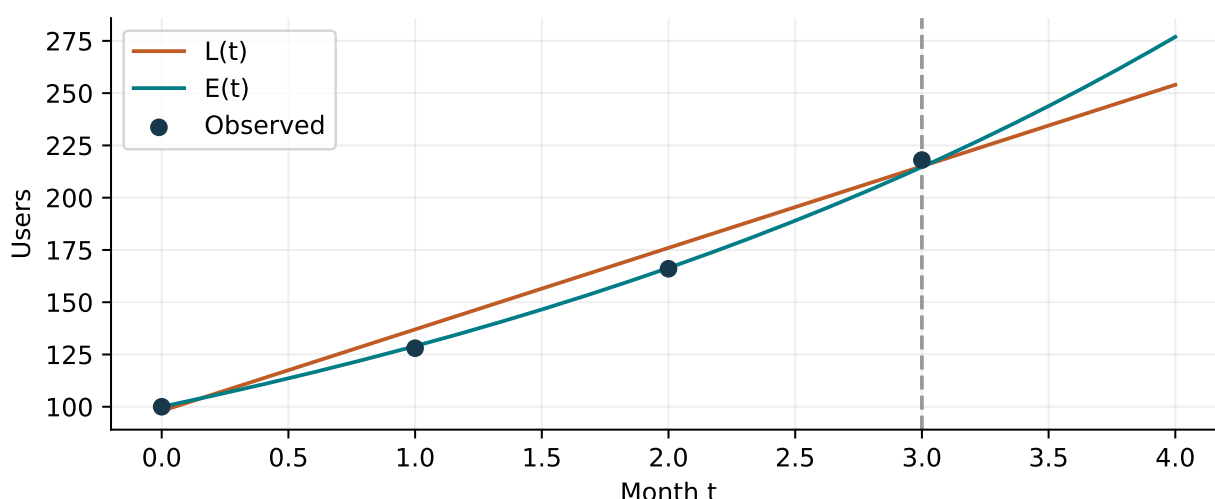


図 3：点は観測値、橙は一次、青緑は指数モデル。破線より右は観測範囲の外です。

選んだ指数モデルの 4 か月後の予測は、

$$E(4) = 100(1.29)^4 = 276.922881 \approx \boxed{277人}.$$

小さい RSS は、将来の正しさの証明ではない

ここで比べたのは指定された 2 つのモデルの、4 時点での当てはまりです。最適なモデルを一般に決定したわけではありません。予測には、新しい時点のデータでの検証と、増加が続く理由の確認が必要です。

8 追加問題と解答

問 A：単位と傾き

費用 $C(q) = 800 + 150q$ 円について、固定費、1 個あたりの変動費、4 個追加したときの増加額を求めよ。

解答：固定費 800 円、変動費 150 円/個、増加額 $150 \cdot 4 = \boxed{600}$ 円。

問 B：範囲内での最小化

$f(x) = (x - 4)^2 + 2$ の $0 \leq x \leq 3$ における最小値と、そのときの x を求めよ。

解答：頂点 4 は範囲外なので、最も近い端点 $x = \boxed{3}$ で最小値 $\boxed{3}$ です。

問 C：指数減衰

初めに 500 個あり、每期 20% ずつ減る。 t 期後の個数のモデルと、3 期後の値を求めよ。

解答： $N_t = 500(0.8)^t$ 、 $N_3 = 500(0.512) = \boxed{256}$ 。

問 D：倍加時間

$N(t) = 80 \cdot 2^{t/3}$ の倍加時間と 6 時点後の値を求めよ。

解答： t が 3 増えると 2 倍なので、倍加時間は $\boxed{3}$ 。 $N(6) = 80 \cdot 4 = \boxed{320}$ 。

問 E：周期モデルの読み取り

$y = 20 + 5 \sin(\pi(t - 2)/6)$ の振幅、周期、最大値、 $2 \leq t < 14$ で最大となる時点を求めよ。

解答：振幅 $\boxed{5}$ 、周期 $\boxed{12}$ 、最大値 $\boxed{25}$ 。 $\pi(t - 2)/6 = \pi/2$ より $t = \boxed{5}$ 。

問 F：平均を代入すれば期待値か

X が 0 と 2 を確率 $1/2$ ずつ取るとき、 $E[X^2]$ と $(E[X])^2$ を比較せよ。

解答： $E[X^2] = (0 + 4)/2 = \boxed{2}$ 、 $(E[X])^2 = 1^2 = \boxed{1}$ 。非線形関数では一般に一致しません。

問 G：データの差と比

等間隔データ 3, 6, 12, 24 について、隣同士の差と比を求め、一次モデルと指数モデルのどちらが自然か答えよ。

解答：差は 3, 6, 12、比は 2, 2, 2。一定なのは比なので、時点を 0 から数えるなら $y_t = 3 \cdot 2^t$ という指数モデルがこのデータに一致します。