数理計画法　第10回

第3章　非線形計画

3.2　制約なし最適化

担当:　塩浦昭義
(情報科学研究科　助教授)
shioura@dais.is.tohoku.ac.jp

復習：制約なし最適化問題，勾配ベクトル

入力：目的関数 f(x) のみ

$$\boxed{\text{最小化}\quad f(x)\quad \text{条件}\quad \text{なし}\quad (x \in R^n)}$$

関数 f の勾配ベクトル

$$\nabla f(\boldsymbol{x}) = \begin{pmatrix} \dfrac{\partial f}{\partial x_1} \\ \vdots \\ \dfrac{\partial f}{\partial x_n} \end{pmatrix}$$

一変数関数の場合は

$$\nabla f(x) = f'(x)$$

復習：一次のテイラー展開[p.89]

関数 f は勾配ベクトルを傾きとする**線形関数により近似**できる

関数 f(x) の x=a における一次のテイラー展開
（a は定数ベクトル）

$$\boxed{f(x) = \underbrace{f(a) + \nabla f(a)^T (x-a)} + \varphi(x-a)}$$

この部分が線形関数, f(x) を近似

関数 $\varphi(x-a) = \varphi(x_1 - a_1, x_2 - a_2, \ldots, x_n - a_n)$ は
$x_1 - a_1, x_2 - a_2, \ldots, x_n - a_n$ に関する
２次以上の項からなる n 変数多項式関数
(定数項，一次の項は全く含まれない)

復習:一次のテイラー展開

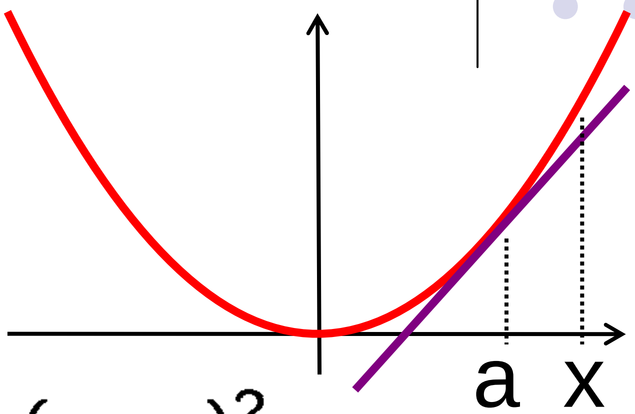
例1 : $f_1(x) = x^2 \qquad f_1'(x) = 2x,$

f_1 の x=a における一次のテイラー展開

$$f(x) = a^2 + 2a(x-a) + \varphi(x-a)$$

ここで

$$\varphi(x-a) = f(x) - \{a^2 + 2a(x-a)\} = (x-a)^2$$



例2 : $f_2(x) = \log x \qquad f_2'(x) = 1/x$

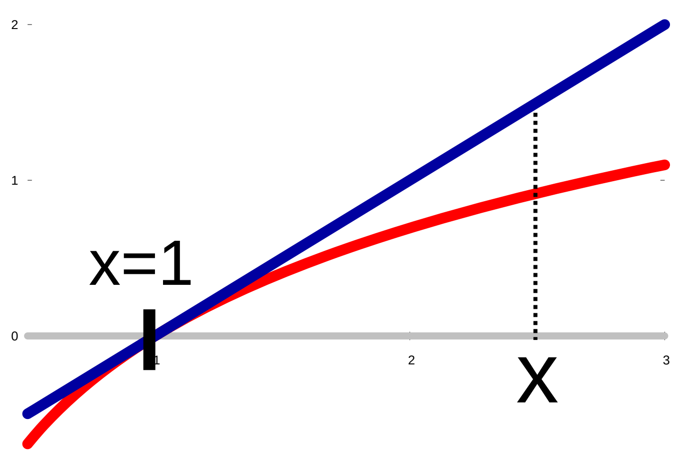
f_2 の x=1 における一次のテイラー展開

$$f_2(x) = 0 + \frac{1}{1}(x-1) + \varphi(x-1)$$

ここで

$$\varphi(x-1)$$

$$= -\frac{(x-1)^2}{2} + \frac{(x-1)^3}{3} - \frac{(x-1)^4}{4} + \cdots$$



勾配ベクトルの性質

勾配ベクトルと逆の方向に進むと関数値が減る

性質: 任意のベクトル $y \in \mathbf{R}^n$ に対し, $\nabla f(y) \neq 0$ ならば
十分小さい $\delta > 0$ に対して $f(y - \delta \nabla f(y)) < f(y)$

証明: $d = -\delta \nabla f(y)$ とおく（δ: 正の実数)

一次のテイラー展開において x = y + d, a = y とおくと,

$$f(y+d) = f(y) + \nabla f(y)^T d + \varphi(d)$$
$$= f(y) - \delta ||\nabla f(y)||^2 + \varphi(-\delta \nabla f(y))$$

φは2次以上の項からなる多項式関数

$\Rightarrow$ $\varphi(-\delta \nabla f(y))$ は δ に関する2次以上の項からなる
一変数関数

勾配ベクトルの性質

勾配ベクトルと逆の方向に進むと関数値が減る

性質: 任意のベクトル $y \in \mathbf{R}^n$ に対し, $\nabla f(y) \neq 0$ ならば
十分小さい $\delta > 0$ に対して $f(y - \delta \nabla f(y)) < f(y)$

証明の続き:

$\varphi(-\delta \nabla f(y))$ は δ に関する2次以上の項からなる一変数関数

$\Rightarrow \varphi(-\delta \nabla f(y))/\delta$ は δ に関する1次以上の項からなる

$\therefore$ δ を十分小さくすると, $\varphi(-\delta \nabla f(y))/\delta$ は0 に近づく

$$\therefore\ -\delta \|\nabla f(y)\|^2 + \varphi(-\delta \nabla f(y))$$
$$= -\delta \{\|\nabla f(y)\|^2 - \varphi(-\delta \nabla f(y))/\delta\} < 0$$

$$\therefore f(y - \delta \nabla f(y)) < f(y)$$

復習:最適性条件

制約なし最適化問題:　最小化　$f(x)$

最適性条件:ベクトル x が非線形計画問題の最適解であるための必要条件

定義: x は停留点 ⇔ $\nabla f(x) = 0$

> 定理(制約なし最適化問題の最適性条件):
> x^*: 制約なし問題の最適解 ⇒ x^*は停留点

最適性条件 [p.97]

定理(制約なし最適化問題の最適性条件):

x^*: 制約なし問題の最適解 ⇒ $\nabla f(x^*) = 0$

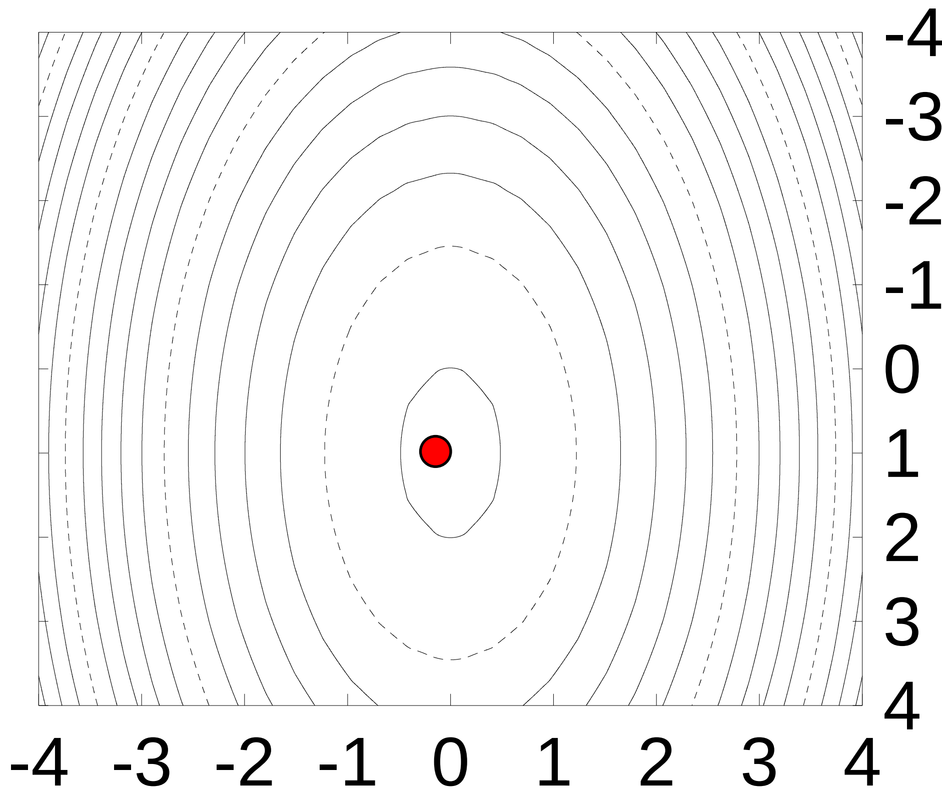
例: $f(\boldsymbol{x}) = x_1^2 - 2x_1 + 4x_2^2$

$$= (x_1 - 1)^2 + 4x_2^2 - 1$$

$(x_1, x_2) = (1,0)$ が最適解

$$\nabla f(\boldsymbol{x}) = \begin{pmatrix} 2x_1 - 2 \\ 8x_2 \end{pmatrix}$$

$$\nabla f(1,0) = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$



最適性条件 [p.97]

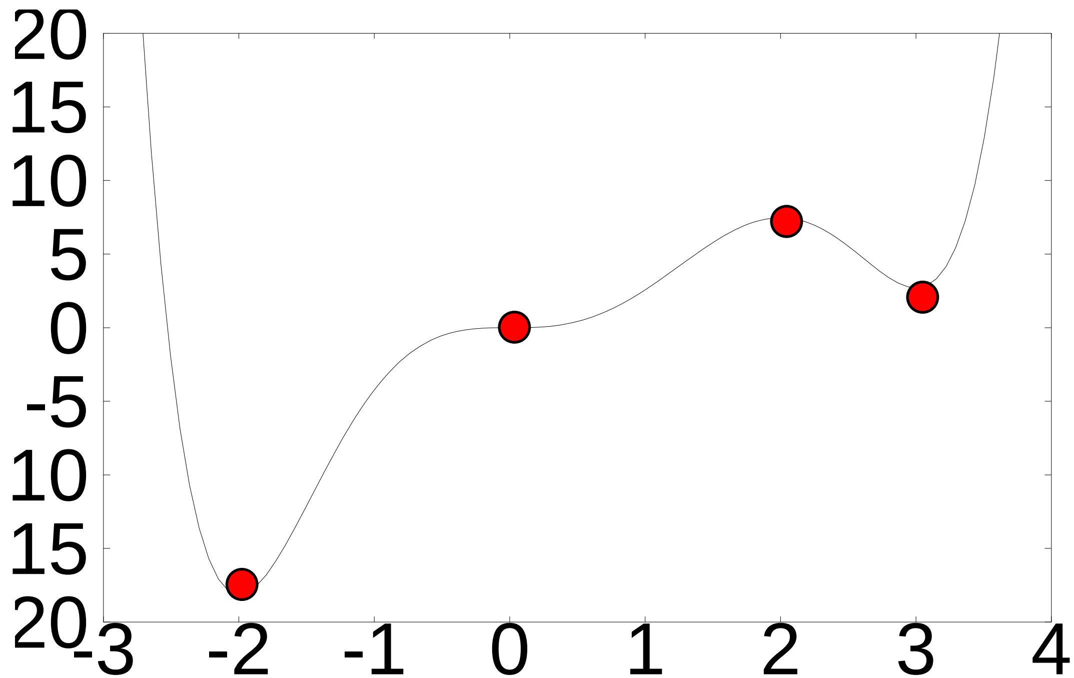
※「x*は停留点 ⇒ x*は最適解」は必ずしも成り立たない

例：

$$f(x) = \frac{1}{6}x^6 - \frac{3}{5}x^5 - x^4 + 4x^3$$

$$\nabla f(x) = (x+2)x^2(x-2)(x-3)$$

停留点はx = -2, 0, 2, 3
最適解はx = -2 のみ



極小解、極大解、鞍点 [p.99]

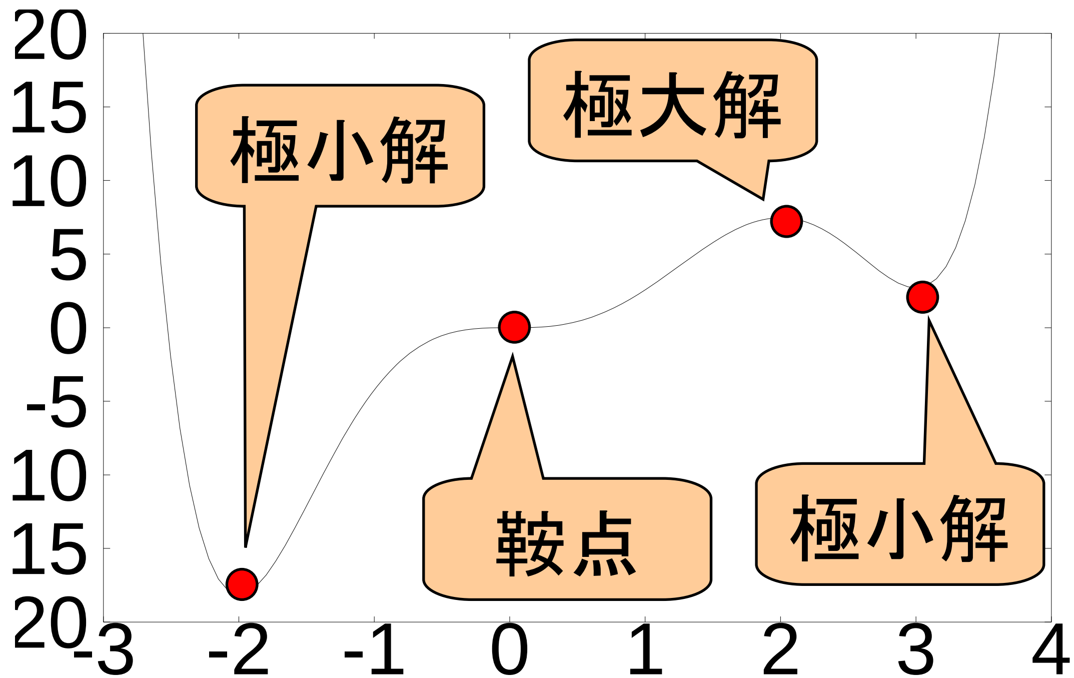
停留点 x^* の分類

極小解: x^* の付近だけに注目したとき、x^* は最小

> あるδ>0 が存在して、$\|x - x^*\| \leqq \delta$ を満たす
> すべての x に対して $f(x) \geqq f(x^*)$

極大解: x^* の付近だけに注目したとき、x^* は最大

鞍点 : 極小点でも
極大点でもない停留点



制約なし問題の解法1:最急降下法 [p.102]

最急降下法のアイディア:
勾配ベクトルと逆の方向に進むと関数値が減る

現在の点 x を x－ $\alpha \nabla f(x)$ により更新

⇒ 関数値 f(x) を減らしていく

ステップサイズ

ステップサイズの選び方:
次の一変数最適化問題を(近似的に)解く

最小化 $f(x - \alpha \nabla f(x))$ 条件 $\alpha > 0$

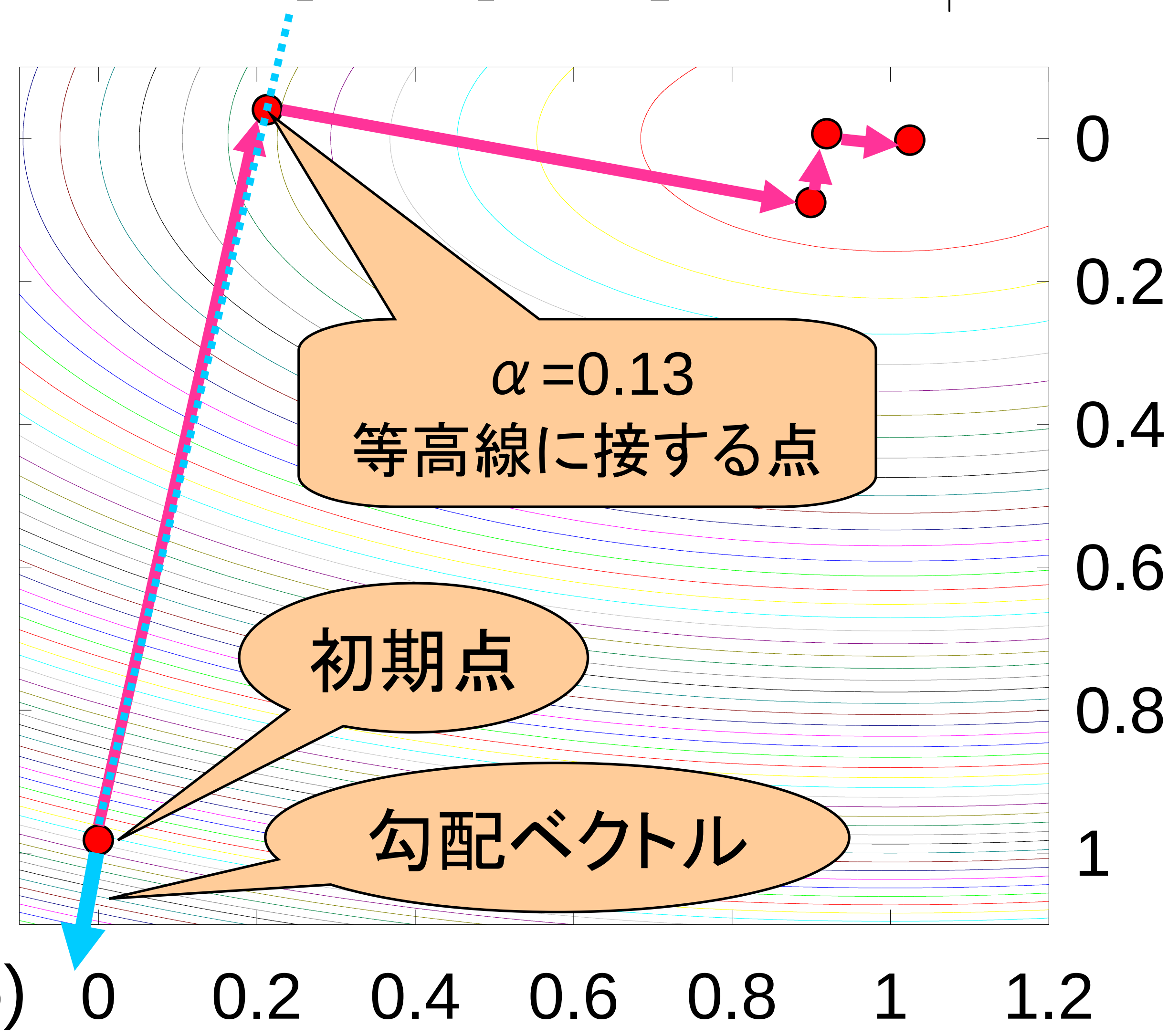
直線探索と呼ばれる

最急降下法の実行例 [p.103]

教科書の例3.2,3.8: $f(\boldsymbol{x}) = x_1^2 - 2x_1 + 4x_2^2$

$$\nabla f(\boldsymbol{x}) = \begin{pmatrix} 2x_1 - 2 \\ 8x_2 \end{pmatrix}$$

- $(x_1,x_2) = (0,1)$ からスタート
- $\nabla f(0,1) = (-2,8)$
- $f(0 + 2\alpha, 1 - 8\alpha)$ を最小にするのは $\alpha=0.13$
- 次の点は $(x_1,x_2) = (0.26,-0.05)$



最急降下法のアルゴリズム [p.102]

入力：関数 f とその勾配ベクトル ∇f

初期点 x^0

ステップ0： $k=0$ とする

ステップ1： x^k が最適解に十分近ければ終了

ステップ2：最急降下方向 $-\nabla f(x^k)$ を計算

ステップ3：直線探索問題

最小化　$f(x^k - \alpha \nabla f(x^k))$　条件　$\alpha > 0$

を解き、解を α^k とする

ステップ4： $x^{k+1} = x^k - \alpha^k \nabla f(x^k)$ とおく

ステップ5： $k=k+1$ として、ステップ1に戻る

ヘッセ行列 [p.90]

関数 f のヘッセ行列

$$\mathrm{H}\, f(\boldsymbol{x}) = \begin{pmatrix} \dfrac{\partial^2 f}{\partial x_1 \partial x_1} & \dfrac{\partial^2 f}{\partial x_1 \partial x_2} & \cdots & \dfrac{\partial^2 f}{\partial x_1 \partial x_n} \\ \dfrac{\partial^2 f}{\partial x_2 \partial x_1} & \dfrac{\partial^2 f}{\partial x_2 \partial x_2} & \cdots & \dfrac{\partial^2 f}{\partial x_2 \partial x_n} \\ \vdots & \vdots & \ddots & \vdots \\ \dfrac{\partial^2 f}{\partial x_n \partial x_1} & \dfrac{\partial^2 f}{\partial x_n \partial x_2} & \cdots & \dfrac{\partial^2 f}{\partial x_n \partial x_n} \end{pmatrix}$$

一変数関数の場合は2階導関数に一致

$$\mathrm{H}\, f(x) = f''(x)$$

ヘッセ行列（続き）[p.89]

例：

$$f_1(x) = x^2 \qquad \nabla f_1(x) = f_1'(x) = 2x,\ \mathrm{H} f_1(x) = f_1''(x) = 2$$

$$f_2(x_1,\ x_2) = \sin x_1 + \cos x_2 \Longrightarrow \nabla f_2(\boldsymbol{x}) = \begin{pmatrix} \cos x_1 \\ -\sin x_2 \end{pmatrix}$$

$$\Longrightarrow \mathrm{H} f_2(\boldsymbol{x}) = \begin{pmatrix} -\sin x_1 & 0 \\ 0 & -\cos x_2 \end{pmatrix}$$

$$f_3(x_1,\ x_2) = x_1 x_2 + \log x_1$$

$$\Longrightarrow \nabla f_3(\boldsymbol{x}) = \begin{pmatrix} x_2 + 1/x_1 \\ x_1 \end{pmatrix} \Longrightarrow \mathrm{H} f_3(\boldsymbol{x}) = \begin{pmatrix} -1/x_1^2 & 1 \\ 1 & 0 \end{pmatrix}$$

レポート問題

問題１:下記の4つの関数のヘッセ行列を計算しなさい

$$f_1(x_1,\ x_2)\ =\ x_1\ +\ 2x_2$$

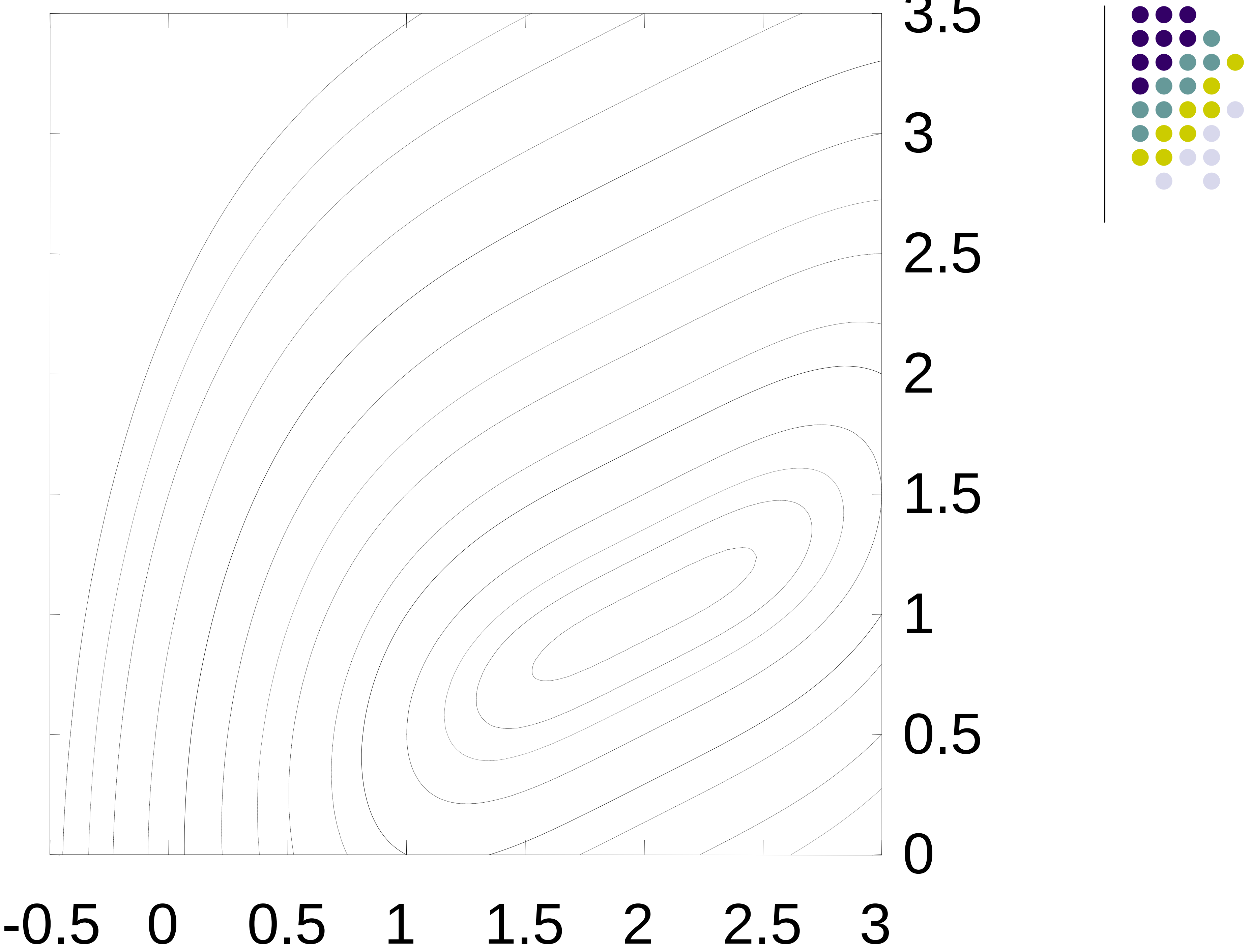
$$f_2(x_1,\ x_2)\ =\ x_1^2\ + x_2^2\ -1$$

$$f_3(x_1,\ x_2)\ =\ x_1\log x_2 - x_2\log x_1$$

(ただし，$\boldsymbol{x}$はn次元ベクトル，$\boldsymbol{V}$は$n\times n$対称行列)

問題3：関数 f(x,y) = $(x – 2)^4 + (x – 2y)^2$ に対して、初期点を(0, 3) として最急降下法を適用せよ。資料に添付してある等高線の図を使って実行すること.（数値はおおまかに計算すればよい）

ポイント: 点の動きを表す折れ線の角度は必ず90度


3.5
3
2.5
2
1.5
1
0.5
0
-0.5
0
0.5
1
1.5
2
2.5
3