計量とポテンシャル勾配による力学系のタイプ

林@ JAIST 知識科学

平成 12 年 3 月 9 日

1 QR アルゴリズム型 ($m = 1$)

有限非周期 Toda 方程式 (Moser 力学系)

$$\frac{dP_i}{dt} = \frac{d^2Q_i}{dt^2} = e^{-(Q_i - Q_{i-1})} - e^{-(Q_{i+1} - Q_i)}$$

ここで, $a_i \stackrel{\text{def}}{=} \frac{1}{2}e^{-\frac{Q_{i+1}-Q_i}{2}}$, $b_i \stackrel{\text{def}}{=} \frac{P_i}{2}$ より,

$$\frac{da_i}{dt} = a_i(b_{i+1} - b_i),$$
$$\frac{db_i}{dt} = 2(a_i^2 - a_{i-1}^2)$$

Moser の作用角変数を用いて式 (1) に変換.

一般化された n 種の Lotka-Volterra 系 式 (1) において, $y^j \stackrel{\text{def}}{=} (x^j)^2$ とすると Lotka-Volterra 系

$$\frac{dy^j}{dt} = 2x^j\frac{dx^j}{dt} = y^j(-2\lambda_j + \sum_k 2\lambda_k y^k)$$

が得られる. ここで, $-2\lambda_j$ は個体 j の内的減少 (増加) 率 r_j で, $2\lambda_k$ は個体 k から j への影響度 a_{jk} に相当する.

ニューロンの Hebb 学習 結合重みの更新/学習（出力 $z = \sum_i s_i x_i$）:

$$\frac{ds_j}{dt} = -\gamma z^2 s_j + \gamma z x_j$$

入力 $x_i(t)$ で上式を平均化して, $E[X^T X]$ を対角化する行列 G を用い, $diag\Lambda \overset{\mathrm{def}}{=} \gamma G E[X^T X] G^T$, $R \overset{\mathrm{def}}{=} GE[S]$ とおいて, 式 (1) に変換. 但し, 勾配系の右辺やポテンシャルの符合が負になる.

QR アルゴリズム 行列の QR 分解： $M = QR$, $M' = Q^T M Q = Q'R'$ の反復により, $M^{(\infty)} = (QQ'Q'' \ldots)^T MQQ'Q'' \ldots$ から固有値の対角行列を求める.
行列の指数関数 $\exp(tL_0) = Q(t)R(t)$ に対する, $L(t) = Q(t)^T L_0 Q(t)$ の Lax 形式の $t = 0 \sim 1$ が QR の 1 ステップに一致する.

以上をまとめると次のような勾配系が得られる。

勾配系（計量なし)：

$$\frac{dx^j}{dt} = \pm x^j \left(-\lambda_j + \sum_k \lambda_k (x^k)^2 \right) = -\frac{\partial V_Q(x)}{\partial x^j} \tag{1}$$

ポテンシャル（分子は Replicator 方程式の平均適応度に対応)：

$$V_Q(x) \overset{\mathrm{def}}{=} \pm \frac{1}{2} \frac{\sum_k \lambda_k (x^k)^2}{\sum_k (x^k)^2}$$

1. 非線形変数変換 (線形化できるが, 計量を介した双対座標ではない)：

$$(x^j)^2 \overset{\mathrm{def}}{=} \frac{e^{y_j}}{\sum_k e^{y_k}}, \quad \sum_j (x^j)^2 = 1 : \ trivial$$

$$\frac{dy_j}{dt} = -2\lambda_j$$

2. Explicit な解表示：

$$\frac{dx^j}{dt} = x^j \left(-\lambda_j + \sum_k \lambda_k (x^k)^2 \right)$$

に対して, $y^j \overset{\mathrm{def}}{=} \log x^j$, $q(x) \overset{\mathrm{def}}{=} \sum_k \lambda_k (x^k)^2$ とおくと

$$\frac{dy^j}{dt} = \frac{1}{x^j} \frac{dx^j}{dt} = -\lambda_j + q(x)$$

より, $y^j(t) = -\lambda_j t + \int_{s=0}^t q(s) ds + const.$ が得られる. 不変量 $\sum_j (x^j)^2 = 1$ を用いてこれを整理すると, 以下の Explicit な解表示が得られる.

$$x^j(t) = \frac{x^j(0) \exp(-\lambda_j t)}{\left(\sum_k x^k(0) \exp(-\lambda_k t) \right)^{1/2}}$$

2　単体上線形計画-内点法型 ($m=2$)

$$\min \quad \sum_j c_j x_j$$

$$s.t. \quad \sum_j x_j = 1, \quad x_j \geq 0$$

数理進化の Shahshahani 勾配系（Replicator 方程式）の 1 種：

$$\frac{dx^i}{dt} = -c_i(x^i)^2 + x^i \sum_k c_k (x^k)^2 = -\sum_j g^{ij} \frac{\partial V_L(x)}{\partial x^j} \tag{2}$$

初期値が内点なら, $\sum_j \frac{dx^j}{dt} = 0$ で内点性を保持（単体内に平衡点あり）.
ポテンシャル（平均適応度 $m_{ij} = c_i \delta_i^j$ に相当）：

$$V_L(x) \stackrel{\text{def}}{=} \frac{1}{2} \frac{\sum_k c_k (x^k)^2}{(\sum_k x^k)^2}$$

単体上の Shahshahani 計量：

$$g_{ij} = \partial_i \partial_j \psi(x) = \frac{1}{x^i} \delta_i^j,$$

$$\psi(x) \stackrel{\text{def}}{=} \sum_j x^i \log x^i$$

Legendre 変換（一般に y における直線／ ∇^*-測地線ではない)：

$$y_i \stackrel{\text{def}}{=} \log x^i = \partial_i \psi(x), \quad \varphi(y) \stackrel{\text{def}}{=} \sum_i x^i y_i - \psi(x)$$

$$\frac{dy_i}{dt} = \sum_j \frac{\partial y_i}{\partial x^j} \frac{dx^j}{dt} = -\frac{\partial V_L(x)}{\partial x^i} = c_i x^i - \sum_k c_k (x^k)^2$$

3　Moser-Karmarkar 型 ($m \neq 1, 3$)

$m \neq 1, 3$ を実パラメータとするより一般的な力学系の族：

$$\frac{dx^i}{dt} = -c_i (x^i)^m - x^i \sum_k c_k (x^k)^2$$

これは, 不変多様体 $P_m \stackrel{\rm def}{=} \{x \geq 0 | \ \sum_i (x^i)^{3-m} = 1\}$ 上の計量 $g_{ij} \stackrel{\rm def}{=} (x^i)^{1-m}\delta_i^j$ と, (最小化される) ポテンシャル関数

$$U(x) \stackrel{\rm def}{=} \frac{1}{2}\sum_k c_k (x^k)^2 \left(\sum_i (x^i)^{3-m}\right)^{\frac{2}{m-3}}$$

を用いた以下の勾配系と等価である.

$$\frac{dx^i}{dt} = -\sum_j g^{ij}\frac{\partial U}{\partial x^i}$$

また, 変数変換 $z^i \stackrel{\rm def}{=} (x^i)^{3-m}$ を施せば, 計量 g_{ij} は z^i に関する単体上の Shahshahani 計量となる.

4 Affine-Projective 変換-内点法型

Affine Scaling Trajectory：標準 LP 問題

$$\min \quad < c, x >$$

$$s.t. \quad \begin{cases} Ax = b, \\ x \geq 0 \end{cases}$$

$$\frac{dx}{dt} = -X\pi_{(AX)^\perp}(Xc) = -\nabla_G | < c, x > |_F$$

ここで, $X \stackrel{\rm def}{=} diag\{x^1, x^2, \ldots, x^n\}$, $\pi_{(AX)^\perp}$ は AX の null 空間への射影である. $\nabla_G | < c, x > |_F$ は, $x > 0$ の log-barrier 関数で定義される R_n^+ の計量 $G = X^{-2}$(scaling) に関する $x^T G x = const.$ のもとで, $F = \{x : \ Ax = b\}$ における目的関数 $< c, x >$ を最小化する方向ベクトルである.

Projective Scaling Trajectory: 正準 LP 問題

(解析的中心を通れば $p_{cen} = 0$)

$$\min \quad < c, x >$$

$$s.t. \quad \begin{cases} Ax = 0, \\ < e, x > = n, \\ x \geq 0 \end{cases}$$

$$\frac{dx}{dt} = -X\pi_{(AX)^\perp}(Xc) + \frac{1}{n} < Xe, \pi_{(AX)^\perp}(Xc) > Xe = p_{aff} + p_{cen}$$

5　べき乗法型

Rayleigh 商による最大固有値（線形差分）

$$\max \quad R_D(r) \stackrel{\text{def}}{=} \frac{< r, Dr >}{||r||^2}$$

$$s.t. \;\; ||r||^2 = \sum_i r_i^2 = 1$$

$R_D(r)$ の勾配系 $\frac{dr}{dt} = Dr- < Dr, r > r$ を, $Q^T DQ = diag(\lambda_i, \ldots, \lambda_n)$, $z = Q^T r$, $x^i = z_i^2$ にて式（3）に変換.

Canonical-form LP(無制約の場合)

$$\min \quad \sum_j c_j x_j$$

$$s.t. \;\; \sum_j x_j = 1, \quad x_j \geq 0,$$

制約条件 $\sum_j x_j = 1$ を満たす内点で, 負エントロピー $\psi = \frac{1}{2}\sum_j x_j \log x_j$ の Hessian を計量 $g_{ij} = \frac{1}{2x_i}\delta_i^j$ とする, $V_P(x) \stackrel{\text{def}}{=} \frac{\sum_k c_k x_k}{\sum_j x_j}$ の勾配系：

$$\frac{dx^i}{dt} = -\sum_j g^{ij} \frac{\partial V_P(x)}{\partial x^j} = -2c_i x^i + x^i \sum_k 2c_k x^k$$

$\Rightarrow$ 差分（線形レベルでの Euler 差分）化して, 制約条件 $\sum_j A_{ij} x_j = 0$ を満たすように修正すると, 反復ごとに射影行列を計算する必要なし.

勾配系（制約上の最急上昇／下降）：

$$\frac{dx^i}{dt} = -\lambda_i x^i + x^i \sum_k \lambda_k x^k = -\sum_j g^{ij} \frac{\partial V_P(x)}{\partial x^j} \qquad (3)$$

初期値が内点なら, $\sum_j \frac{dx^j}{dt} = 0$ で内点性を保持.

ポテンシャル：

$$V_P(x) \stackrel{\text{def}}{=} \frac{\sum_k \lambda_k x^k}{\sum_k x^k}$$

Legendre 変換（$\partial_i y_j = \partial_i \partial_j \psi(x) = g_{ij}$ だが, 測地線とは限らない）：

$$y_i \stackrel{\text{def}}{=} \log x^i = \partial_i \psi(x), \quad \varphi(y) \stackrel{\text{def}}{=} \sum_i x^i y_i - \psi(x)$$

$$\frac{dy_i}{dt} = \sum_j \frac{\partial y_i}{\partial x^j} \frac{dx^j}{dt} = -\frac{\partial V_P(x)}{\partial x^i} = \lambda_i - \sum_k \lambda_k x^k$$

6 凸計画-内点法型

6.1 Affine-scaling 法

不等式制約の線形計画問題

$$\min \quad \sum_j c_j x^j \stackrel{\mathrm{def}}{=} f(x)$$

$$s.t. \quad \sum_{j=1}^{n} A_{ij} x^j \le b_i \quad (i = 1, 2, \ldots, m)$$

Barrier 関数： $U(x) \stackrel{\mathrm{def}}{=} -\sum_i \log(b_i - \sum_j A_{ij} x^j)$
不等式制約に対する計量：

$$g_{ij} = \partial_i y_j = \partial_i \partial_j U(x) = \sum_k \frac{A_{ki} A_{kj}}{(b_k - \sum_l A_{kl} x^l)^2}$$

$$y_j = \partial_j U(x) = \sum_k \frac{A_{kj}}{b_k - \sum_l A_{kl} x^l}$$

x 座標における Projective Scaling 法は, $< c, x >= c_{opt}$ を ∞ に移した変換座標 $y = \Psi(x)$ における log-barrier 関数の Newton 法と等価.

凸計画問題

$$\min \quad \sum_j c_j x^j \stackrel{\mathrm{def}}{=} f(x)$$

$$s.t. \ \ x \in M \subset R^m$$

Self-concordant barrier 関数： $\Psi_t(x) \stackrel{\mathrm{def}}{=} t \times c^t x + \psi(x)$
最適条件： $\partial_i \Psi_t(x) = 0 \Leftrightarrow y_i = -c_i t \quad (y_i \stackrel{\mathrm{def}}{=} \partial_i \psi(x)$ なので)
凸領域 M の計量： $g_{ij} = \partial_i y_j = \partial_i \partial_j \Psi(x) = \partial_i \partial_j \psi(x)$
$\Rightarrow \Psi_t(x)$ の最小化が主パス追跡法で, Newton 法で多項式計算時間.
式 (4) の軌跡は, 初期値が解析的中心ならば中心曲線に一致.

勾配系（Barrier 関数に対する計量を持つ affine-scaling 軌跡）：

$$\frac{dx^j}{dt} = -\sum_j g^{ij} c_j = -\sum_j g^{ij} \frac{\partial f(x)}{\partial x^j} \tag{4}$$

$$\frac{df(x)}{dt} = \frac{d \sum_i c_i x^i}{dt} = \sum_i c_i \frac{dx^i}{dt} = -\sum_{ij} g^{ij} c_i c_j < 0$$

Legendre 変換（線形化でき, 線形時間の直線 y は ∇^*-測地線）：

$$\frac{dy_i}{dt} = \sum_j \frac{\partial y_i}{\partial x^j}\frac{dx^j}{dt} = \sum_j g_{ij}\frac{dx^j}{dt} = -c_i$$

$$y_i \overset{\text{def}}{=} \partial_i \psi(x), \quad \varphi(y) \overset{\text{def}}{=} \sum_i x^i y_i - \psi(x)$$

6.2 半正定値問題

線形システムの安定状態フィードバックゲイン行列の集合も, $PD(n)$ 中でパラメータ化できる (同じ数理構造を持つ).

$$\min \quad \sum_j c_j x^j \overset{\text{def}}{=} f(x)$$

$$s.t. \;\; P(x) \overset{\text{def}}{=} E_0 + \sum_{i=1}^m x^i E_i \geq O$$

基底の対称行列： $E_i \in Sym(n)$, $(i = 0, 1, 2, \ldots, m)$, $P(x)$ の許容領域：$\mathcal{L} = PD(n) \bigcap (E_0 + span\{E_i\}_{i=1}^m)$

Self-concordant barrier 関数： $\psi(x) \overset{\text{def}}{=} -\log\det P(x)$ を考え, 計量 $g_{ij}(x) = \partial_i\partial_j\psi(x)$, 接続係数 $\Gamma_{ki,j} = 0$, $\Gamma^*_{kj,i} = \partial_k\partial_i\partial_j\psi(x) = \partial_k g_{ij}(x)$ から情報幾何構造 $(\mathcal{L}, g, \nabla, \nabla^*)$ を定める.

$PD(n)$ 全体では, 主双対座標 θ, η によって陽に表現できる.

$$P(x) = \sum_i \theta^i E_i = \left(\sum_i \eta_i E^i\right)^{-1}$$

$$-tr\{E^i E_j\} = \delta_i^j$$

ゆえに, ∇-測地線上では P が線形, ∇^*-測地線上では P^{-1} が線形.

Legendre 変換：

$$\eta_i = -tr\left(E_i P^{-1}\right), \quad \theta^i = -tr\left(E^i P\right),$$

$$\eta_i \overset{\text{def}}{=} \partial_i\psi, \quad \theta^i \overset{\text{def}}{=} \partial^i\varphi, \quad \varphi(\eta) \overset{\text{def}}{=} \sum_i \theta^i\eta_i - \psi(\theta)$$

一方, 凸錐 $PD(n)$ の部分領域 $\mathcal{L}$ では, 上記の陽表示は一般に（二重自己平行でない時）は不可. すなわち, ∇^*-測地線（affine-scaling 軌跡）に対する双対座標 η（の直線）から主座標 θ への逆 Legendre 変換を行なう, Newton 反復計算が必要となる.

二重自己平行でない時の勾配系：

$$\frac{d\theta}{dt} = -\begin{pmatrix} G_1^{-1}c \\ 0, \end{pmatrix} \Leftrightarrow \begin{cases} \theta^i = x^i + \theta_0^i & i = 1, \ldots, m, \\ \theta^k = \theta_0^k & k = m+1, \ldots, n, \end{cases}$$

$$\frac{d\eta}{dt} = -\begin{pmatrix} c \\ G_2^T G_1^{-1}c, \end{pmatrix} \Leftrightarrow \begin{cases} \eta_i = y_i & i = 1, \ldots, m, \\ \eta_k = \frac{\partial\psi(\theta)}{\partial\theta^k}|_\theta & k = m+1, \ldots, n, \end{cases}$$

$$G = \left(\frac{\partial^2\psi}{\partial\theta^i\partial\theta^j}\right) = \begin{pmatrix} G_1 & G_2 \\ G_2^T & G_3 \end{pmatrix}$$

$\mathcal{L}$ 上の主双対座標系 x, y と, $PD(n)$ の座標系 θ, η の関係（θ_0 は E_0 の θ 座標）. ここで, $\psi(\theta) = -\log\det\sum_i \theta^i E_i$ である.

7 ニュートン法型（情報幾何）

双対平坦な $S = \{p(x;\theta)\}$ 上の ∇^*-測地線を表す勾配系：

$$\frac{d\theta^i}{dt} = -\sum_j g^{ij}\frac{\partial D}{\partial\theta^j} = -\sum_j \left[\frac{\partial^2 D}{\partial\theta^i\partial\theta^j}\right]^{-1}\frac{\partial D}{\partial\theta^j} \tag{5}$$

ダイバージェンス（Contrast 汎関数)：

$$D(p||q) \stackrel{\text{def}}{=} \psi(\theta(p)) + \varphi(\eta(q)) - \sum_i \theta^i(p)\eta_i(q)$$

S 上の Fisher 計量が D の Hessian に一致：

$$g_{ij} \stackrel{\text{def}}{=} \sum_x \partial_i l(x;\theta)\partial_j l(x;\theta)p(x;\theta) = \partial_i\partial_j\psi(\theta) = \partial_i\partial_j D,$$

Legendre 変換（線形化でき, 指数時間で漸近収束する直線)：

$$\frac{d\eta_i}{dt} = \sum_j \frac{\partial y_i}{\partial x^j}\frac{dx^j}{dt} = -\frac{\partial D}{\partial\theta^i} = -(\eta_i - \eta_i(q))$$

$$\eta_i \stackrel{\text{def}}{=} \partial_i\psi(\theta), \quad \varphi(y) \stackrel{\text{def}}{=} \sum_i x^i y_i - \psi(x)$$

8　外点法的中心化ニュートン法型

設計方針： 最適解が満たすべき Kuhn-Tucker 条件の中の等式部（中心平坦化写像 $\Phi(x)=0$）に, ニュートン法を適用して探索方向ベクトルを得る. これは, 双対ギャップを零にする LP の主双対内点法に通じる.

$\Rightarrow$ 制約充足 $g(x)=0$ と目的関数の極小条件 $\nabla f(x)=0$ の和をニュートン法で解く.

$$\min f(x_1, x_2, \ldots, x_n)$$
$$s.t \begin{cases} g_1(x_1, x_2, \ldots, x_n) = 0 \\ \vdots \\ g_m(x_1, x_2, \ldots, x_n) = 0 \end{cases} \quad (m \le n)$$

Lagrange 関数最適化

$$L(x,\lambda) \stackrel{\text{def}}{=} f(x) + \lambda^T g(x)$$

この Lagrange 関数 $L(x,\lambda)$ の最適条件は, 勾配 $\nabla f(x)$ と Jacobi 行列 $J(x) \stackrel{\text{def}}{=} [\partial g_i/\partial x_j]$ を用いて,

Lagrange 条件： $\nabla_x L = \nabla f(x) + J^T(x)\lambda = 0$

許容条件： $g(x) = 0$

なので, これらの等式を Newton-Raphson 法で解く.

$$\begin{bmatrix} H(x,\lambda) & J^T(x) \\ J(x) & 0 \end{bmatrix} \begin{pmatrix} \Delta x \\ \lambda^+ \end{pmatrix} = -\begin{pmatrix} \nabla f(x) \\ g(x) \end{pmatrix}$$

ここで, $\lambda^+ = \lambda + \Delta\lambda$ である. また, $H(x,\lambda)$ は $L(x,\lambda)$ の Hessian で, これを設計行列 $N(x)$ で置き換えるとその選び方によって種々の反復解法が得られる. 但し, $N(x)$ は $kerJ(x)$ への射影行列 $P(x)$ に対して, $P^T NP \ge 0$, $rankP^T NP = n - m$ を満たす対称行列とする.

CGPM（連続化射影勾配法） 勾配の $kerJ$ 上への射影

$$\Phi(x) \stackrel{\text{def}}{=} -(I - J^+(x)J(x))\nabla f(x) = -P(x)\nabla f(x)$$

が零になるように,

$$\frac{dx}{dt} = \Phi(x) = 0$$

を解く．ここで, $P(x)$ は射影行列, $J^+(x)$ は $J(x)$ の Moore-Penrose 一般化逆行列とする．よって, 以下の非線形連立方程式を得る．右辺の行列が計量 G に相当する．

$$\begin{bmatrix} I & J^T(x) \\ J(x) & 0 \end{bmatrix} \begin{bmatrix} \Phi(x) \\ \Lambda(x) \end{bmatrix} = - \begin{bmatrix} \nabla f(x) \\ 0 \end{bmatrix}$$

これは, $\Lambda \stackrel{\mathrm{def}}{=} (J^+)^T \nabla f = (JJ^T)^{-1} J \nabla f$ と,

$$J^T \Lambda = J^T (JJ^T)^{-1} J \nabla f = J^+ J \nabla f,$$

$$\Phi + J^T \Lambda = (I - J^+ J) \nabla f + J^+ J \nabla f = \nabla f,$$

$$J\Phi = J(I - J^+ J)\nabla f = (J - JJ^+ J)\nabla f = 0$$

より確かめられる．

FIGAPM（制約を強いる修正項 $J^+ g$ を付ける）

$$\Psi(x) \stackrel{\mathrm{def}}{=} -(I - J^+(x)J(x))\nabla f(x) - J^+(x) g(x)$$

が零になるように,

$$\frac{dx}{dt} = \Phi(x) = 0$$

を解く．ここで,

$$\frac{dg(x)}{dt} = \sum_i \frac{\partial g}{\partial x_i} \frac{dx_i}{dt} = J(x)\Phi(x)$$

であり, $\frac{dg}{dt} = 0$ の代わりに制約を強いるように, $(g_1, \ldots, g_m)$ の座標系における直線上の軌跡を表す

$$\frac{dg}{dt} - g$$

を用いる（制約充足すれば $g(x) = 0$ よりこれらは同値）．

よって, 以下の非線形連立方程式を得る．

$$\begin{bmatrix} I & J^T(x) \\ J(x) & 0 \end{bmatrix} \begin{bmatrix} \Psi(x) \\ \bar{\Lambda}(x) \end{bmatrix} = - \begin{bmatrix} \nabla f(x) \\ g(x) \end{bmatrix}$$

ここで, $\bar{\Lambda} = (JJ^T)^{-1}(J\nabla f + g)$ である．

$\Rightarrow$ 不等式系（ニュートンベクトルと中心化ベクトルの和）にも拡張できる．