

graph-based SLAM の解説

千葉工業大学 上田隆一

2017 年 4 月 18 日

1 はじめに

この文章は、[1] などのチュートリアルを見ても数式の細かいところ分からない graph-based SLAM について、実際の計算方法を細かく解説するためのものである。実装例は GitHub の `ryuichiueda/probrobo_practice` の `./graph-based_SLAM/graph-based_slam.ipynb` にあり、この文章はこの実装の数学的な解説となる。ただし、コードで使った記号と本文章の記号は一部異なる。

2 問題

平面上を移動し、向きを持ち、カメラでランドマーク観測ができるロボットで graph-based SLAM を実行する方法を考える。ランドマークは環境にいくつか存在し、ロボットからは互いに識別でき、距離と見える方角が観測できる。また、2 つの観測がどの方角から観測されたものか、相対的に分かるものとする。

2.1 ロボットの姿勢と座標系

世界座標系 Σ_w におけるロボットの姿勢（位置と向き）を

$$\boldsymbol{x} = \begin{bmatrix} x \\ y \\ \theta \end{bmatrix} \quad (1)$$

で表す。また、 $[x \ y]^T$ を原点として、 X 軸が世界座標系で θ の方向を向いているロボット座標系 Σ_r を考える。これらの関係を図 1 に示す。

離散的な時刻 $t = 0, 1, 2, \dots, T$ を考える。時刻の集合を \mathcal{T} で表す。時刻 t における世界座標系でのロボットの姿勢の真値を \boldsymbol{x}_t^* で表す。ロボットはデッドレコニングで \boldsymbol{x}_t の推定値 $\hat{\boldsymbol{x}}_t$ を認識するが、ロボットの動作は雑音の影響を受けるため、 \boldsymbol{x}_t^* と $\hat{\boldsymbol{x}}_t$ の間には誤差が発生する。

ロボットは一つの行動ごとに $\hat{\boldsymbol{x}}_t$ を記録していく。全時刻の推定姿勢を

$$\hat{\boldsymbol{x}}_{0:T} = \{\hat{\boldsymbol{x}}_0, \hat{\boldsymbol{x}}_1, \dots, \hat{\boldsymbol{x}}_T\} \quad (2)$$

と表すこととする。

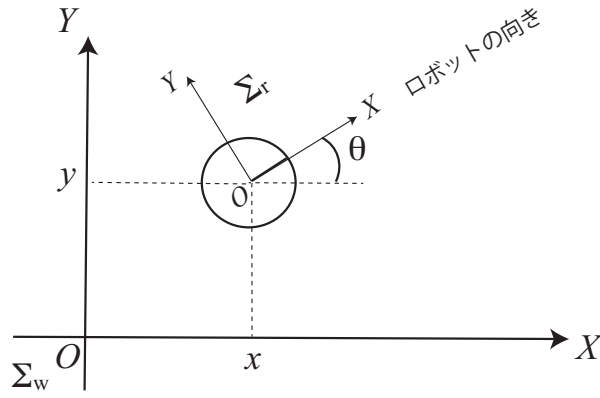


図1 世界座標系とロボットの姿勢

2.2 観測

環境中にいくつかランドマークが存在していると仮定する。時刻 t におけるロボット座標系 Σ_r を Σ_{rt} と表すこととすると、ロボットには、時刻 t において、全ランドマークのうちいくつかを計測する。

2.2.1 ランドマークの識別

ロボットからは、一度観測したランドマークは、後の時刻で観測したときに、どのランドマークか識別できることとする。ロボットは観測したランドマークに ID を与えて管理することにする。ID は c と表し（番号でも文字列でもなんでも良い）、ID として c を与えられたランドマークを L_c と表す。ロボットが認識しているランドマークの ID の集合を C で表す。

2.2.2 ランドマークの姿勢計測

ロボットは Σ_{rt} においてランドマーク L_c を観測したとき、 L_c までの距離 $d_{c,t}$ と、ランドマークが見える方向 $\varphi_{c,t}$ を計測値として得る。また、ランドマークも方角を持ち（ロボットの θ に相当）、ロボットに対してどの方角を向いているか分かることとする。この値を $\psi_{c,t}$ とする*1。図2にこれらの記号の関係を示す。

2.2.3 計測値の記録

ロボットが時刻 t で得るランドマーク全ての計測値の集合は、 $Z_t = \{z_{c,t} = (d_{c,t}, \varphi_{c,t}, \psi_{c,t}) | c \in C, c: \text{観測したランドマークの ID}\}$ で表すことができ、これもロボットは各時刻ごとに記録する。 Z_t の集合を $Z_{0:T}$ で表す。

2.2.4 計測値の誤差

ランドマーク観測から得られる計測値には正規分布に従う雑音が入ると仮定する。また、ロボットは雑音の傾向を知っていることとする。

*1 この仮定は実用上強すぎるが、実際には、後の計算式から分かるように、2つの姿勢間での値 $\psi_{c,t}, \psi_{c,t'}$ の差だけが分かれば良い。例えば、2点間で得られた画像の向きを画像処理から割り出すなどの処理で、この差は得られる。

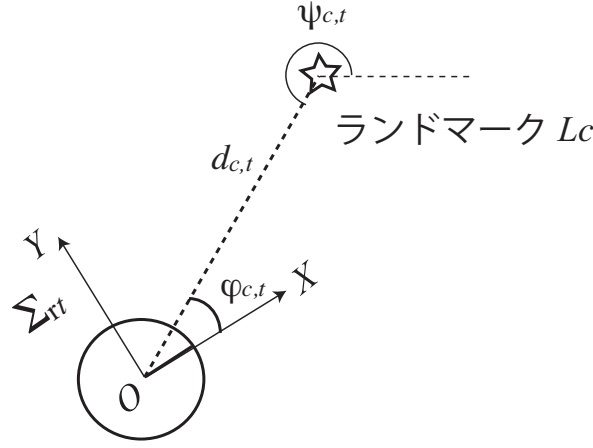


図2 計測値

$d_{c,t}$ には、10[%] の雑音が入る^{*2}。正確には真値を $d_{c,t}^*$ とすると、

$$d_{c,t} \sim \mathcal{N}(d_{c,t}^*, (d_{c,t}^*/10)^2) \quad (3)$$

にしたがって $d_{c,t}$ が生成される。ここで $\mathcal{N}(\mu, \sigma^2)$ は平均値 μ 、標準偏差 σ の正規分布を表す。

また、 $\varphi_{c,t}, \psi_{c,t}$ にも $3\pi/180[\text{rad}]$ の標準偏差で雑音が入る。これらの真値をそれぞれ $\varphi_{c,t}^*, \psi_{c,t}^*$ とすると、 $\varphi_{c,t}, \psi_{c,t}$ は、

$$\varphi_{c,t} \sim \mathcal{N}(\varphi_{c,t}^*, (3\pi/180)^2) \quad (4)$$

$$\psi_{c,t} \sim \mathcal{N}(\psi_{c,t}^*, (3\pi/180)^2) \quad (5)$$

で生成される。

2.3 完全 SLAM 問題

ここで、 $Z_{0:T}$ 、デッドレコニングでの計測値 $\hat{x}_{0:T}$ から、真値 $x_{0:T}^*$ を推定する問題を考える。この問題を、任意のロボットの軌跡 $x_{0:T}$ がどれだけ $Z_{0:T}$ を説明するかを評価関数として定式化し、この評価関数を用いた最適化問題として定義する。この評価関数については、次章の実装の中で説明する。

3 graph-based SLAM の実装例

実装の一例を示す。

3.1 グラフのエッジを作る

問題を解くために、次のノード（頂点）、エッジ（辺、アーク）を持つグラフを考える。ノードはロボットの各時刻の姿勢に対応し、エッジは、同じランドマークを計測した2つのノード間に設けられる。

^{*2}「10[%]」は変数にすべきだが、記号が増えて理解の妨げになるので固定値として説明する。

各ノードは、ロボットの各時刻の推定姿勢 $\mathbf{x}_{0:T}$ を変数として有する。これらの推定姿勢は計算の途中で変化していくが、その初期値はデッドレコニングから得られた $\hat{\mathbf{x}}_{0:T}$ とする。

エッジは、両端の2つのノードについて、2つのランドマーク観測から計算できる2つのノードの相対姿勢を表すベクトル $\boldsymbol{\mu}_{c,t,t'}$ その正確さを情報として持つ。ベクトル $\boldsymbol{\mu}_{c,t,t'}$ は、時刻 t の姿勢から時刻 t' の姿勢への向きを正とする。また、「正確さ」というのは世界座標系での情報行列 $\Omega_{c,t,t'}$ で表現される。情報行列とは共分散行列の逆行列で、 $\Omega_{c,t,t'}$ を計算する際の共分散行列 $\Sigma_{c,t,t'}$ は、2.2.4 項でモデル化した雑音から計算する。計算方法は後述する。

このようなグラフを作ると、1つのエッジと両端のノードは、

- エッジの持つ $\boldsymbol{\mu}_{c,t,t'}$
- 両ノードの持つ推定姿勢の差 $\mathbf{x}_{t'} - \mathbf{x}_t$

の、2つの重複した相対姿勢の情報を持つことになる。この1つのエッジと両端のノードに話を限ると、推定姿勢 $\mathbf{x}_{t'}, \mathbf{x}_t$ を変化させて、次の誤差

$$\mathbf{e}_{c,t,t'} = \mathbf{x}_{t'} - \mathbf{x}_t - \boldsymbol{\mu}_{c,t,t'} \quad (6)$$

をゼロにすると、推定姿勢がランドマーク観測の計測値を最も説明するものになる。一方、ノードは他のエッジとも繋がっているので、こちらだけをゼロに合わせると別のエッジでの誤差が増えるため、全体として最適な推定姿勢を探さなければならない。

3.1.1 $\boldsymbol{\mu}_{c,t,t'}, \mathbf{e}_{c,t,t'}$ の計算

2つの時刻 t, t' で同じランドマーク L_c を観測したときのロボットとランドマークの幾何的な関係を図3に示す。二つの観測を図の点線の矢印のように世界座標系にベクトルとして描くと、時刻 t の観測のベクトルに、時刻 t' の観測のベクトルの向きを反転して足せば $\boldsymbol{\mu}_{c,t,t'}$ が求まるのが分かる。

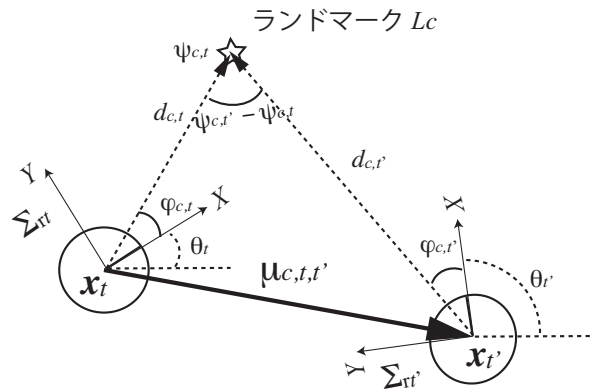


図3 ランドマークの計測値から2点の相対姿勢を求める

したがって、 $\boldsymbol{\mu}_{c,t,t'}$ (x, y, θ 座標のベクトルとなる) は次のような単純な幾何計算で求まる*3。

$$\boldsymbol{\mu}_{c,t,t'} = \begin{bmatrix} d_{c,t} \cos(\theta_t + \varphi_{c,t}) \\ d_{c,t} \sin(\theta_t + \varphi_{c,t}) \\ \psi_{c,t} \end{bmatrix} - \begin{bmatrix} d_{c,t'} \cos(\theta_{t'} + \varphi_{c,t'}) \\ d_{c,t'} \sin(\theta_{t'} + \varphi_{c,t'}) \\ \psi_{c,t'} \end{bmatrix} \quad (7)$$

*3 おそらく ψ は θ で置き換えられるので ψ を使わない実装もできるが、まだ自分自身では検証していない。

また、式 (6) から、

$$\begin{aligned} \mathbf{e}_{c,t,t'} &= \mathbf{x}_{t'} - \mathbf{x}_t - \begin{bmatrix} d_{c,t} \cos(\theta_t + \varphi_{c,t}) \\ d_{c,t} \sin(\theta_t + \varphi_{c,t}) \\ \psi_{c,t} \end{bmatrix} + \begin{bmatrix} d_{c,t'} \cos(\theta_{t'} + \varphi_{c,t'}) \\ d_{c,t'} \sin(\theta_{t'} + \varphi_{c,t'}) \\ \psi_{c,t'} \end{bmatrix} \\ &= \begin{bmatrix} x_{t'} - x_t - d_{c,t} \cos(\theta_t + \varphi_{c,t}) + d_{c,t'} \cos(\theta_{t'} + \varphi_{c,t'}) \\ y_{t'} - y_t - d_{c,t} \sin(\theta_t + \varphi_{c,t}) + d_{c,t'} \sin(\theta_{t'} + \varphi_{c,t'}) \\ \theta_{t'} - \theta_t - \psi_{c,t} + \psi_{c,t'} \end{bmatrix} \end{aligned} \quad (8)$$

となる。

3.1.2 $\Sigma_{c,t,t'}, \Omega_{c,t,t'}$ の計算

共分散行列 $\Sigma_{c,t,t'}$ は、 t, t' におけるランドマーク L_c の計測の誤差に関する共分散行列 $\Sigma_{c,t}, \Sigma_{c,t'}$ をそれぞれの計測座標系で求め、世界座標系に座標変換（回転）して足し合わせたものとなる。ここで言う「計測座標系」とは、ロボットからランドマークが計測された向きに引いた線を X 軸とする座標系である。図 4 に、ロボット座標系、計測座標系の関係と、ランドマークの計測値の誤差に関する共分散行列を示す。

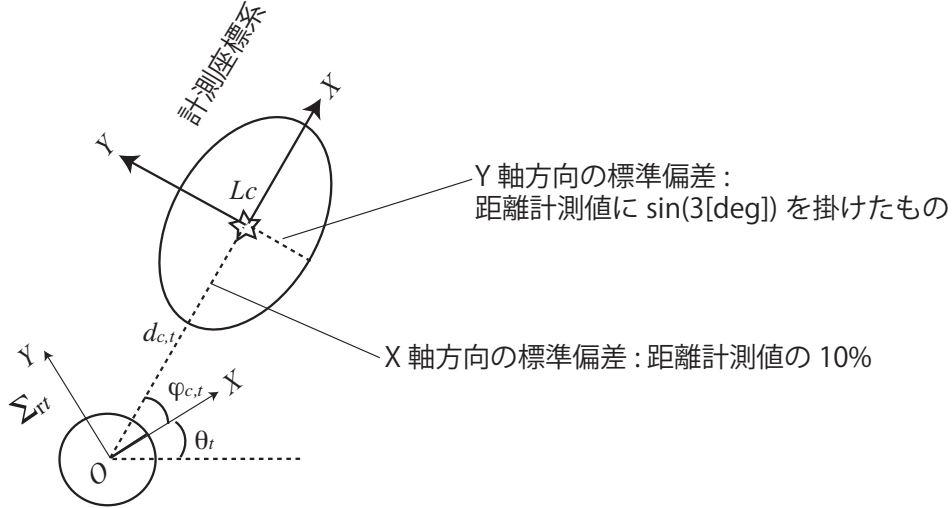


図 4 ランドマークの計測値の不確かさを表す共分散行列

時刻 t での L_c の計測に関する雑音を計測値周りのガウス分布で表すと、計測座標系におけるこの分布の共分散行列 $\Sigma_{c,t}$ は、次のようになる。

$$\Sigma_{c,t} = \begin{bmatrix} (d_{c,t}/10)^2 & 0 & 0 \\ 0 & [(d_{c,t}/10) \sin(3\pi/180)]^2 & 0 \\ 0 & 0 & (3\pi/180)^2 \cdot 2 \end{bmatrix} \quad (9)$$

対角成分はそれぞれ計測座標系での X, Y 軸、 θ 方向の共分散であり、式 (3-5) の分布の分散に由来する。 Y 軸の共分散は、ランドマークの向きの計測誤差に、距離の計測値を掛けたものになる*4。また、 θ 方向の共分散については、 ϕ, ψ の計測の雑音両方が不確かさの原因になるので、分散の値は一方のものの倍になる。

*4 小さい角度なので、 $\sin(3\pi/180)$ は $3\pi/180$ に近似しても良い。

次に、計測座標系で計算した $\Sigma_{c,t}$ を世界座標系の誤差楕円に変換する。デッドレコニングの値、ランドマークの計測値から、計測座標系は世界座標系から $\theta_{t'} + \varphi_{c,t'}$ だけ回転しているとみなすことができるので、 $\Sigma_{c,t}$ をこの角度だけ回転してやればよい。この回転行列は、

$$R_{c,t} = \begin{bmatrix} c & s & 0 \\ -s & c & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad (10)$$

となる。ここで、

$$\begin{aligned} c &= \cos(\theta_t + \varphi_{c,t}) \\ s &= \sin(\theta_t + \varphi_{c,t}) \end{aligned}$$

である。この回転行列を $\Sigma_{c,t}, \Sigma_{c,t'}$ に対してそれぞれ求めると、

$$\Sigma_{c,t,t'} = R_{c,t} \Sigma_{c,t} R_{c,t}^T + R_{c,t'} \Sigma_{c,t'} R_{c,t'}^T \quad (11)$$

$$\Omega_{c,t,t'} = \Sigma_{c,t,t'}^{-1} \quad (12)$$

となる。この時点で $\Sigma_{c,t,t'}$ の各要素は、全て既知の値から計算された実数になっているはずである。

3.2 最適化問題を作る

2.3 節で棚上げした問題定義を行う。まず、ロボットがオンラインで得たデータから得られる全通りの $\mathbf{e}_{c,t,t'}$ で集合 \mathcal{E} を作る。このとき、 \mathcal{E} の中の各要素がゼロベクトルになるように $\mathbf{x}_{0:T}$ を $\boldsymbol{\mu}_{c,t,t'}$ に近づけていくと、ランドマーク観測で得られた計測値を説明するロボットの軌跡 $\mathbf{x}_{0:T}$ が得られる。しかし、ランドマーク観測で得られた計測値には雑音が含まれ、全ての計測値と一致するような軌跡は得られない。また、2.2.4 項の定義では、遠いところでランドマークを観測した時の距離の計測値 d と、近いところで計測した時の d では、近いところで計測したものの方が精度が良くなる。そのため、遠いところの計測値と近いところの計測値が互いに矛盾しているときは、軌跡の推定の際は近いところの計測値を優先すべきである。

3.2.1 マハラノビス距離

以上を勘案する上で、情報行列 $\Omega_{c,t,t'}$ 、ベクトル $\mathbf{e}_{c,t,t'}$ に対して次のような値を考えることは有用である。

$$f_{\Omega_{c,t,t'}}(\mathbf{x}_t, \mathbf{x}_{t'}) = \mathbf{e}_{c,t,t'}^T \Omega_{c,t,t'} \mathbf{e}_{c,t,t'} \quad (13)$$

この値はマハラノビス距離を二乗した値である。マハラノビス距離は、誤差のベクトル \mathbf{e} の二乗誤差を計算するときに、ただ内積 $\mathbf{e}^T \mathbf{e}$ を計算するのではなく、間に情報行列を入れて重みをつけたものと解釈できる。1次元で考えると、同じ大きさの誤差に対し、共分散が大きいと（つまり誤差が大きいと予想される場合）マハラノビス距離が小さくなり、共分散が小さいとマハラノビス距離が大きくなる。

3.2.2 最適化する式

各ベクトル $\mathbf{e}_{c,t,t'}$ に対して、マハラノビス距離の二乗の和:

$$F(\mathbf{x}_{0:T}) = \sum_{\mathbf{e}_{c,t,t'} \in \mathcal{E}} \mathbf{e}_{c,t,t'}^T \Omega_{c,t,t'} \mathbf{e}_{c,t,t'} \quad (14)$$

を最小にする最適化問題を考える。このようにマハラノビス距離を用いることで、ランドマークの計測値が信頼できる（この場合では近いところでランドマークを観測したときの計測値で構成される）エッジの影響が、信頼できないエッジよりも誤差に対して F の変化に大きな影響を与えるため、姿勢を推定するときに優先される。

3.3 $\mathbf{e}_{c,t,t'}$ の勾配を求める

\mathbf{x}_t と $\mathbf{x}_{t'}$ を変化させたときに $\mathbf{e}_{c,t,t'}$ がどれだけ変化するかを求める。まず、 \mathbf{x}_t と $\mathbf{x}_{t'}$ を連結したべ 6 次元のベクトル

$$\begin{aligned} \mathbf{x}_{tt'} &= [\mathbf{x}_t \ \mathbf{x}_{t'}]^T \\ &= [x_t \ y_t \ \theta_t \ x_{t'} \ y_{t'} \ \theta_{t'}]^T \end{aligned}$$

を考える。これを $\Delta \mathbf{x}$ だけ微小変化させたときの $\mathbf{e}_{c,t,t'}$ ($\mathbf{x}_{tt'}$ と同じ変数を持つ 6 次元ベクトル \mathbf{x} の関数とみなす) の変化量は、次の近似で説明できる。

$$\mathbf{e}_{c,t,t'}(\mathbf{x}_{tt'} + \Delta \mathbf{x}) = \mathbf{e}_{c,t,t'}(\mathbf{x}_{tt'}) + \left. \frac{\partial \mathbf{e}_{c,t,t'}}{\partial \mathbf{x}} \right|_{\mathbf{x}=\mathbf{x}_{tt'}} \Delta \mathbf{x} \quad (15)$$

ここでヤコビ行列 $\left. \frac{\partial \mathbf{e}_{c,t,t'}}{\partial \mathbf{x}} \right|_{\mathbf{x}=\mathbf{x}_{tt'}}$ を $J_{c,t,t'}$ と表すと、

$$J_{c,t,t'} = \begin{bmatrix} -1 & 0 & d_{c,t} \sin(\theta_t + \varphi_{c,t}) & 1 & 0 & -d_{c,t'} \sin(\theta_{t'} + \varphi_{c,t'}) \\ 0 & -1 & -d_{c,t} \cos(\theta_t + \varphi_{c,t}) & 0 & 1 & d_{c,t'} \cos(\theta_{t'} + \varphi_{c,t'}) \\ 0 & 0 & -1 & 0 & 0 & 1 \end{bmatrix} \quad (16)$$

となる。後の計算のため、この行列の左半分と右半分を分けて、

$$A_{c,t,t'} = \begin{bmatrix} -1 & 0 & d_{c,t} \sin(\theta_t + \varphi_{c,t}) \\ 0 & -1 & -d_{c,t} \cos(\theta_t + \varphi_{c,t}) \\ 0 & 0 & -1 \end{bmatrix} \quad (17)$$

$$B_{c,t,t'} = \begin{bmatrix} 1 & 0 & -d_{c,t'} \sin(\theta_{t'} + \varphi_{c,t'}) \\ 0 & 1 & d_{c,t'} \cos(\theta_{t'} + \varphi_{c,t'}) \\ 0 & 0 & 1 \end{bmatrix} \quad (18)$$

を定義しておく。この二つの行列は、最適化問題を解く際に、それぞれ $\mathbf{x}_t, \mathbf{x}_{t'}$ をどの方向に動かすと $f_{\Omega_{c,t,t'}}(\mathbf{x}_t, \mathbf{x}_{t'})$ の変化量が大きいかという情報を持っている。

3.4 問題を解く

まず、 $\mathbf{x}_{0:T}$ の各変数を順番に並べた $3(T+1)$ 行 $3(T+1)$ 列の大きな情報行列 H を考える。この情報行列には、各エッジの誤差の情報を足し込んでいく。具体的には、 H の時刻 i に対応する

3 行と、時刻 j に対応する 3 列で構成される 3×3 の部分行列 $H_{[ij]}$ を考えたとき、次のように各エッジのデータを足し込んでいく。

$$H_{[tt]} \leftarrow H_{[tt]} + A_{c,t,t'}^T \Omega_{c,t,t'} A_{c,t,t'} \quad (19)$$

$$H_{[tt']} \leftarrow H_{[tt']} + A_{c,t,t'}^T \Omega_{c,t,t'} B_{c,t,t'} \quad (20)$$

$$H_{[t't]} \leftarrow H_{[t't]} + B_{c,t,t'}^T \Omega_{c,t,t'} A_{c,t,t'} \quad (21)$$

$$H_{[t't']} \leftarrow H_{[t't']} + B_{c,t,t'}^T \Omega_{c,t,t'} B_{c,t,t'} \quad (22)$$

ただし、ここでは各時刻に最低一つのランドマークの計測値が存在することを仮定した。そうでない場合は、ランドマークの計測値が存在しない時刻は行列を構成するときに除外しておく必要がある。

H は、各姿勢の各変数を動かしたときに F にどれだけ影響を与えるかの情報を持っている。ここで、最初の姿勢 \mathbf{x}_0 を世界座標系の原点に固定すると*5、 \mathbf{x}_0 が動いてはいけなないので、

$$H_{[00]} \leftarrow H_{[00]} + \alpha I \quad (\alpha : \text{大きな正の値}) \quad (23)$$

というように、単位行列に大きな数字をかけたもの（例えば、エッジの情報を H に足し込むときの値の桁と比べて 100 倍や 1000 倍くらいの値）を足し込んでおく。これは、 \mathbf{x}_0 を動かすことに対して大きなペナルティーを与えて、 \mathbf{x}_0 を動かないようにする処置である。

次に、 $3(T+1)$ 次元の「情報ベクトル」 \mathbf{b} を準備する。このベクトルに対しても、時刻 t に対応する \mathbf{b} の部分ベクトル $\mathbf{b}_{[t]}$ を考え、ここに各エッジの情報を次のように足しこむ。

$$\mathbf{b}_{[t]} \leftarrow \mathbf{b}_{[t]} + A_{c,t,t'}^T \Omega_{c,t,t'} \mathbf{e}_{c,t,t'} \quad (24)$$

$$\mathbf{b}_{[t']} \leftarrow \mathbf{b}_{[t']} + B_{c,t,t'}^T \Omega_{c,t,t'} \mathbf{e}_{c,t,t'} \quad (25)$$

このとき、ランドマークの計測値を $\mathbf{x}_{0:T}$ がより良く説明するために必要な変化量を、 $\mathbf{x}_{0:T}$ の変数を順番に並べた $3(T+1)$ 次元ベクトル $\Delta \mathbf{x}$ で表すと、

$$\Delta \mathbf{x} = -H^{-1} \mathbf{b} \quad (26)$$

が成り立つ（説明は後で）。

以上の計算を、 $\mathbf{x}_{0:T}$ を $\Delta \mathbf{x}$ だけ動かし、再度 H, \mathbf{b} を計算して $\Delta \mathbf{x}$ を求め、と繰り返す。すると、初期値に使ったデッドレコニングの値 $\hat{\mathbf{x}}_{0:T}$ がそれなりに良ければ、ランドマークの計測値を全体的にバランスよく説明できる姿勢に $\mathbf{x}_{0:T}$ を $\Delta \mathbf{x}$ が収束する。

参考文献

- [1] Grisetti, G., Kümmerle, R., Stachniss, C. and Burgard, W.: A Tutorial on Graph-Based SLAM, *IEEE Intelligent Transportation Systems Magazine*, Vol. 2 (2010), 31–43.

*5 固定しないと世界座標系が決まらない。