VIO from scratch 5-2, 5-3

Monday, April 6, 2020 9:51 AM

SLAM formulation:

Consider a state ξ, and a measurement r_i. because of noise, we have $p(r_i|\xi)$.

When we have multiple observations, they are independent.

$\therefore$ for $r=(r_1, \cdots, r_n)^T$, $p(r|\xi) = \prod_i p(r_i|\xi)$.

If we know the prior for state, based on GPS, etc, then

$$p(\xi|r) = \frac{p(r|\xi)\,p(\xi)}{p(r)}$$

$$\therefore \xi_{MAP} = \arg\max_{\xi} p(\xi|r)$$

$$= \arg\max_{\xi} \prod_i p(r_i|\xi)\,p(\xi)$$

$$= \arg\min_{\xi} \left[-\sum_i \log p(r_i|\xi) - \log p(\xi)\right]$$

assume $p(r_i|\xi) = \mathcal{N}(\mu_i, \Sigma_i)$, $p(\xi) = \mathcal{N}(\mu_\xi, \Sigma_\xi)$

$$\xi_{MAP} = \arg\min_{\xi} \sum_i \| r_i - \mu_i \|^2_{\Sigma_i} + \|\xi - \mu_\xi\|^2_{\Sigma_\xi}$$

$$J^T \Sigma^{-1} J\, \delta\xi = -J^T \Sigma^{-1} r$$

Multivariate Gaussian:

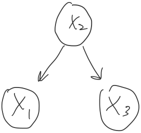
$$p(x) = \frac{1}{z} \exp\left(-\frac{1}{2} x^T \Sigma^{-1} x\right), \quad \Lambda = \Sigma^{-1}.$$

Λ is the information matrix, i.e. weights.

Toy eg 1:

X_2 = temperature outdoor

X_1, X_3 = temperature at room 1,3



$$X_2 = V_2$$
$$X_1 = W_1 X_1 + V_1$$
$$X_3 = W_3 X_2 + V_3$$

V_i are independent Gaussian with σ_i^2

$$\Sigma_{11} = E(X_1 X_1) = E((W_1 V_2 + V_1)(W_1 V_2 + V_1))$$
$$= W_1^2 E(V_2^2) + 2W_1 E(V_1 V_2) + E(V_1^2)$$

$$= w_1^2 \sigma_2^2 + \sigma_1^2$$

$$\Sigma_{22} = \sigma_2^2, \quad \Sigma_{33} = w_3^2 \sigma_2^2 + \sigma_3^2$$

$$\Sigma_{12} = E(x_1 x_2) = E((w_1 v_2 + v_1) v_2) = w_1 \sigma_2^2$$

$$\Sigma_{13} = E((w_1 v_2 + v_1)(w_3 v_2 + v_3)) = w_1 w_3 \sigma_2^2$$

$$\therefore \Sigma = \begin{bmatrix} w_1^2 \sigma_2^2 + \sigma_1^2 & w_1 \sigma_2^2 & w_1 w_3 \sigma_2^2 \\ w_1 \sigma_2^2 & \sigma_2^2 & w_3 \sigma_2^2 \\ w_1 w_3 \sigma_2^2 & w_3 \sigma_2^2 & w_3^2 \sigma_2^2 + \sigma_3^2 \end{bmatrix}$$

$$P(x_1, x_2, x_3) = P(x_2) P(x_1 | x_2) P(x_3 | x_2)$$

$$= \frac{1}{Z_2} \exp\left(-\frac{x_2^2}{2\sigma_2^2}\right) \frac{1}{Z_1} \exp\left(-\frac{(x_1 - w_1 x_2)^2}{2\sigma_1^2}\right) \frac{1}{Z_3} \exp\left(-\frac{(x_3 - w_3 x_2)^2}{2\sigma_3^2}\right)$$

$$P(x_1, x_2, x_3)$$

$$= \frac{1}{Z} \exp\left(-\frac{x_2^2}{2\sigma_2^2} - \frac{(x_1 - w_1 x_2)^2}{2\sigma_1^2} - \frac{(x_3 - w_3 x_2)^2}{2\sigma_3^2}\right)$$

$$= \frac{1}{Z} \exp\left(-x_2^2 \left[\frac{1}{2\sigma_2^2} + \frac{w_1^2}{2\sigma_1^2} - \frac{w_3^2}{2\sigma_3^2}\right] - x_1^2 \frac{1}{2\sigma_1^2} + 2x_1 x_2 \frac{w_1}{2\sigma_1^2} - x_3^2 \frac{1}{2\sigma_3^2} + 2x_3 x_2 \frac{w_3}{2\sigma_3^2}\right)$$

$$= \frac{1}{Z} \exp\left(-\frac{1}{2} [x_1 \; x_2 \; x_3]\right) \underbrace{\begin{bmatrix} \frac{1}{\sigma_1^2} & -\frac{w_1}{\sigma_1^2} & 0 \\ -\frac{w_1}{\sigma_1^2} & \frac{w_1^2}{\sigma_1^2} + \frac{1}{\sigma_2^2} + \frac{w_3^2}{\sigma_1^2} & -\frac{w_3}{\sigma_3^2} \\ 0 & -\frac{w_3}{\sigma_3^2} & \frac{1}{\sigma_3^2} \end{bmatrix}}_{\text{information matrix } \Lambda} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \Bigg)$$

if (i, j) in Λ is 0, then i, j are independent given other variables.
e.g. x_1, x_3 are independent given x_2.

Suppose $w_i > 0$, then if $\Sigma_{ij} > 0$, means two variables are positively correlated.
$\Lambda_{12} < 0$ means given x_3, x_1, x_2 are positively correlated.

toy eg 2:

triangulation: use 2 camera poses to get feature position:

$$x_2 = w_1 x_1 + w_3 x_3 + v_2$$

X_1 X_3

$$\Sigma = \begin{bmatrix} \sigma_1^2 & w_1\sigma_1^2 & \boxed{0} \\ w_1\sigma_1^2 & \sigma_2^2 + w_1^2\sigma_1^2 + w_3^2\sigma_3^2 & w_3\sigma_3^2 \\ \boxed{0} & w_3\sigma_3^2 & \sigma_3^2 \end{bmatrix}$$

$$P(X_1, X_2, X_3) = P(X_1)P(X_3)P(X_2 \mid X_1, X_3)$$

$$= \frac{1}{Z_1}\exp\left(-\frac{X_1^2}{2\sigma_1^2}\right)\frac{1}{Z_3}\exp\left(-\frac{X_3^2}{2\sigma_3^2}\right)\frac{1}{Z_2}\exp\left(-\frac{(X_2 - w_1X_1 - w_3X_3)^2}{2\sigma_2^2}\right)$$

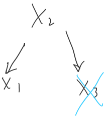
$$= \frac{1}{Z}\exp\left(-\frac{X_1^2}{2\sigma_1^2} - \frac{X_3^2}{2\sigma_3^2} - \frac{(X_2 - w_1X_1 - w_3X_3)^2}{2\sigma_2^2}\right)$$

$$= \frac{1}{Z}\exp\left(-\frac{1}{2}[X_1\ X_2\ X_3]\underbrace{\begin{bmatrix} \frac{1}{\sigma_1^2} + \frac{w_1^2}{\sigma_2^2} & -\frac{w_1}{\sigma_2^2} & \boxed{\frac{w_1w_3}{\sigma_2^2}} \\ -\frac{w_1}{\sigma_2^2} & \frac{1}{\sigma_2^2} & -\frac{w_3}{\sigma_2^2} \\ \boxed{\frac{w_1w_3}{\sigma_2^2}} & -\frac{w_3}{\sigma_2^2} & \frac{1}{\sigma_3^2} + \frac{w_3^2}{\sigma_2^2} \end{bmatrix}}_{\text{information matrix } \Lambda}\begin{bmatrix} X_1 \\ X_2 \\ X_3 \end{bmatrix}\right)$$

Green terms > 0, meaning given X_2, if X_1 ↑, X_3 ↓ ($w_1, w_3 > 0$). Even if Green terms in Covariance is 0, doesn't mean they are uncorrelated.

toy eg 1 without X_3

$$\Sigma = \begin{bmatrix} w_1^2\sigma_2^2 + \sigma_1^2 & w_1\sigma_2^2 & \cancel{w_1w_3\sigma_2^2} \\ w_1\sigma_2^2 & \sigma_2^2 & \cancel{w_3\sigma_2^2} \\ \cancel{w_1w_3\sigma_2^2} & \cancel{w_3\sigma_2^2} & \cancel{w_3^2\sigma_2^2 + \sigma_3^2} \end{bmatrix}$$



$$\Sigma_2 = \begin{bmatrix} w_1^2\sigma_2^2 + \sigma_1^2 & w_1\sigma_2^2 \\ w_1\sigma_2^2 & \sigma_2^2 \end{bmatrix}$$

$$\Sigma^{-1} = \begin{bmatrix} \frac{1}{\sigma_1^2} & -\frac{w_1}{\sigma_1^2} & \cancel{0} \\ -\frac{w_1}{\sigma_1^2} & \frac{w_1^2}{\sigma_1^2} + \frac{1}{\sigma_2^2} + \cancel{\frac{w_3^2}{\sigma_1^2}} & -\cancel{\frac{w_3}{\sigma_3^2}} \\ \cancel{0} & -\cancel{\frac{w_3}{\sigma_3^2}} & \cancel{\frac{1}{\sigma_3^2}} \end{bmatrix}$$

$$\Sigma_2^{-1} = \begin{bmatrix} \frac{1}{\sigma_1^2} & -\frac{w_1}{\sigma_1^2} \\ -\frac{w_1}{\sigma_1^2} & \frac{w_1^2}{\sigma_1^2} + \frac{1}{\sigma_2^2} \end{bmatrix}$$

We need marginalization and Schur's complement.

Given $M = \begin{bmatrix} A & B \\ C & D \end{bmatrix}$,

if D is invertible, $A - BD^{-1}C$ is the Schur's complement of D wrt M

if A is invertible, $D - CA^{-1}B$ is the Schur's complement of A wrt M.

$$\begin{bmatrix} I & 0 \\ -CA^{-1} & I \end{bmatrix}\begin{bmatrix} A & B \\ C & D \end{bmatrix} = \begin{bmatrix} A & B \\ 0 & \Delta_A \end{bmatrix}$$

$$\begin{bmatrix} A & B \\ C & D \end{bmatrix}\begin{bmatrix} I & -A^{-1}B \\ 0 & I \end{bmatrix} = \begin{bmatrix} A & 0 \\ C & \Delta_A \end{bmatrix}$$

where $\Delta_A = D - CA^{-1}B$.

$$\begin{bmatrix} I & 0 \\ -CA^{-1} & I \end{bmatrix}\begin{bmatrix} A & B \\ C & D \end{bmatrix}\begin{bmatrix} I & -A^{-1}B \\ 0 & I \end{bmatrix} = \overbrace{\begin{bmatrix} A & 0 \\ 0 & \Delta_A \end{bmatrix}}^{\text{diag. matrix}}$$

$$\begin{bmatrix} I & 0 \\ CA^{-1} & I \end{bmatrix}\begin{bmatrix} A & 0 \\ 0 & \Delta_A \end{bmatrix}\begin{bmatrix} I & A^{-1}B \\ 0 & I \end{bmatrix} = \begin{bmatrix} A & B \\ C & D \end{bmatrix}$$

$$\therefore \begin{bmatrix} A & B \\ C & D \end{bmatrix} = \begin{bmatrix} I & 0 \\ CA^{-1} & I \end{bmatrix}\begin{bmatrix} A & 0 \\ 0 & \Delta_A \end{bmatrix}\begin{bmatrix} I & A^{-1}B \\ 0 & I \end{bmatrix}$$

$$\begin{bmatrix} A & B \\ C & D \end{bmatrix}^{-1} = \begin{bmatrix} I & -A^{-1}B \\ 0 & I \end{bmatrix}\begin{bmatrix} A^{-1} & 0 \\ 0 & \Delta_A^{-1} \end{bmatrix}\begin{bmatrix} I & 0 \\ -CA^{-1} & I \end{bmatrix}$$

$$\begin{bmatrix} I & -A^{-1}B \\ 0 & I \end{bmatrix}\begin{bmatrix} I & A^{-1}B \\ 0 & I \end{bmatrix} = I$$

toy eg.

$x = \begin{bmatrix} a \\ b \end{bmatrix}$, $K = \begin{bmatrix} A & C^T \\ C & D \end{bmatrix}$, $A = \mathrm{cov}(a,a)$, $D = \mathrm{cov}(b,b)$,

$C = \mathrm{cov}(a,b)$,

$$P(a,b) = P(a)\,P(b|a) \propto \exp\left(-\frac{1}{2}\begin{bmatrix} a \\ b \end{bmatrix}^T \begin{bmatrix} A & C^T \\ C & D \end{bmatrix}^{-1}\begin{bmatrix} a \\ b \end{bmatrix}\right)$$

$P(a,b)$

$$\propto \exp\left(-\frac{1}{2}\begin{bmatrix} a \\ b \end{bmatrix}^T \begin{bmatrix} A & C^T \\ C & D \end{bmatrix}^{-1}\begin{bmatrix} a \\ b \end{bmatrix}\right)$$

$$\propto \exp\left(-\frac{1}{2}\begin{bmatrix} a \\ b \end{bmatrix}^T \begin{bmatrix} I & -A^{-1}C^T \\ 0 & I \end{bmatrix}\begin{bmatrix} A^{-1} & 0 \\ 0 & \Delta_A^{-1} \end{bmatrix}\begin{bmatrix} I & 0 \\ -CA^{-1} & I \end{bmatrix}\begin{bmatrix} a \\ b \end{bmatrix}\right)$$

$$\propto \exp\left(-\frac{1}{2}\begin{bmatrix} a & b - A^{-1}C^T b \end{bmatrix}\begin{bmatrix} A^{-1} & 0 \\ 0 & \Delta_A^{-1} \end{bmatrix}\begin{bmatrix} a \\ b - CA^{-1}a \end{bmatrix}\right)$$

$$\propto \exp\left(-\tfrac{1}{2}\left(a^TA^{-1}a\right)+(b-A^{-1}C^Tb)^T\Delta_A^{-1}(b-A^{-1}C^Tb)\right)$$

$$\propto \underbrace{\exp\left(-\tfrac{1}{2}a^TA^{-1}a\right)}_{p(a)}\underbrace{\exp\left(-\tfrac{1}{2}(b-A^{-1}C^Tb)^T\Delta_A^{-1}(b-A^{-1}C^Tb)\right)}_{p(b|a)}$$

∴ We can get $p(a)$, $p(b|a)$ from $p(a,b)$

$$p(a)=\int_b p(a,b)$$

$$p(a)\propto \exp\left(-\tfrac{1}{2}a^TA^{-1}a\right)\sim \mathcal{N}(0,A).$$

The cov. of marginal is the corresponding submatrix in the joint.

$$P(b|a)\propto \exp\left(-\tfrac{1}{2}(b-A^{-1}C^Tb)^T\Delta_A^{-1}(b-A^{-1}C^Tb)\right)$$

$P(b|a)\sim \mathcal{N}(A^{-1}C^Tb,\Delta_A)$, Covariance is the Schur's complement of a.

In SLAM, we have $J^T\Sigma^{-1}J\delta x=b$, so we deal with the information matrix, instead of covariance.

Suppose we know the information matrix

$$\begin{bmatrix} A & C^T \\ C & D \end{bmatrix}^{-1}=\begin{bmatrix} \Lambda_{aa} & \Lambda_{ab} \\ \Lambda_{ba} & \Lambda_{bb} \end{bmatrix}$$

$$\begin{bmatrix} A & C^T \\ C & D \end{bmatrix}^{-1}=\begin{bmatrix} A^{-1}+A^{-1}C^T\Delta_A^{-1}CA^{-1} & -A^{-1}C^T\Delta_A^{-1} \\ -\Delta_A^{-1}CA^{-1} & \Delta_A^{-1} \end{bmatrix}\triangleq\begin{bmatrix} \Lambda_{aa} & \Lambda_{ab} \\ \Lambda_{ba} & \Lambda_{bb} \end{bmatrix}$$

$$\therefore \Delta_A^{-1}=\Lambda_{bb}.$$

$$A^{-1}=\Lambda_{aa}-\Lambda_{ab}\Lambda_{bb}^{-1}\Lambda_{ba}$$

We want to marginalize x_3 from $P(x_1,x_2,x_3)$, $P(x_1,x_2)$'s information matrix is

$$K_2^{-1}=\Lambda_{aa}-\Lambda_{ab}\Lambda_{bb}^{-1}\Lambda_{ba}$$

$$=\Lambda_{aa}-\begin{bmatrix} 0 \\ -\frac{w_3}{\sigma_3^2} \end{bmatrix}\sigma_3^2\begin{bmatrix} 0 & -\frac{w_3}{\sigma_3^2} \end{bmatrix}$$

$$=\Lambda_{aa}-\begin{bmatrix} 0 & 0 \\ 0 & \frac{w_3}{\sigma_3^2} \end{bmatrix}$$

$$=\begin{bmatrix} \frac{1}{\sigma_1^2} & -\frac{w_1}{\sigma_1^2} \\ -\frac{w_1}{\sigma_1^2} & \frac{w_1^2}{\sigma_1^2}+\frac{1}{\sigma_2^2} \end{bmatrix}$$

this is equivalent to removing the colored term above.

Summary:

$$p(a,b) = \mathcal{N}\left(\begin{bmatrix}\mu_a\\ \mu_b\end{bmatrix}, \begin{bmatrix}\Sigma_{aa} & \Sigma_{ab}\\ \Sigma_{ba} & \Sigma_{bb}\end{bmatrix}\right)$$

$$= \mathcal{N}^{-1}\left(\begin{bmatrix}\eta_a\\ \eta_b\end{bmatrix}, \begin{bmatrix}\Lambda_{aa} & \Lambda_{ab}\\ \Lambda_{ba} & \Lambda_{bb}\end{bmatrix}\right)$$

	marginal	conditional
	$p(a) = \int p(a,b)\,db$	$p(a \mid b) = p(a,b)/p(b)$
cov.	$\mu = \mu_a$ $\Sigma = \Sigma_{aa}$	$\mu' = \mu_a + \Sigma_{ab}\Sigma_{bb}^{-1}(b-\mu_b)$ $\Sigma' = \Sigma_{aa} - \Sigma_{ab}\Sigma_{bb}^{-1}\Sigma_{ba}$
info.	$\eta = \eta_a - \Lambda_{ab}\Lambda_{bb}^{-1}\eta_b$ $\Lambda = \Lambda_{aa} - \Lambda_{ab}\Lambda_{bb}^{-1}\Lambda_{ba}$ (*)	$\eta' = \eta_a - \Lambda_{ab}b$ $\Lambda' = \Lambda_{aa}$

For blue terms, we need to use Schur complement.

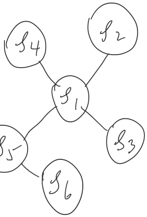
For green terms, no need to use Schur complement

In SLAM, (*) is used the most.

toy eg 3:

$$\xi = \arg\min_{\xi} \frac{1}{2}\sum_i \|r_i\|^2_{\Sigma_i}$$

$$\xi = \begin{bmatrix}\xi_1\\ \xi_2\\ \vdots\\ \xi_6\end{bmatrix}, \quad r = \begin{bmatrix}r_{12}\\ r_{13}\\ r_{14}\\ r_{15}\\ r_{56}\end{bmatrix}$$



circle = vertex = variable to be optimized

edge = residual

Gauss-Newton solution:

$$\underbrace{J^T\Sigma^{-1}J}_{H \text{ or } \Lambda}\,\delta\xi = \underbrace{-J^T\Sigma^{-1}r}_{b}$$

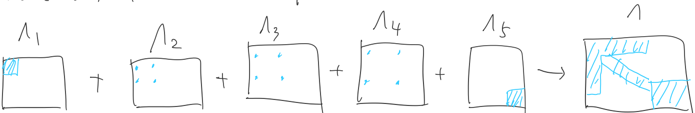
$$J = \frac{\partial r}{\partial \xi} = \begin{bmatrix} \frac{\partial r_{12}}{\partial \xi} \\ \frac{\partial r_{13}}{\partial \xi} \\ \frac{\partial r_{14}}{\partial \xi} \\ \frac{\partial r_{15}}{\partial \xi} \\ \frac{\partial r_{56}}{\partial \xi} \end{bmatrix} = \begin{bmatrix} J_1 \\ J_2 \\ J_3 \\ J_4 \\ J_5 \end{bmatrix}, \quad J^T = [J_1^T \; J_2^T \; J_3^T \; J_4^T \; J_5^T]$$

$$\sum_{i=1}^{5} J_i^T \Sigma_i^{-1} J_i \, \delta\xi = - \sum_{i=1}^{5} J^T \Sigma_i^{-1} r$$

$$J_2 = \frac{\partial r_{13}}{\partial \xi} = \left[\frac{\partial r_{13}}{\partial \xi_1} \quad 0 \quad \frac{\partial r_{13}}{\partial \xi_3} \quad 0 \quad 0 \quad 0 \right]$$

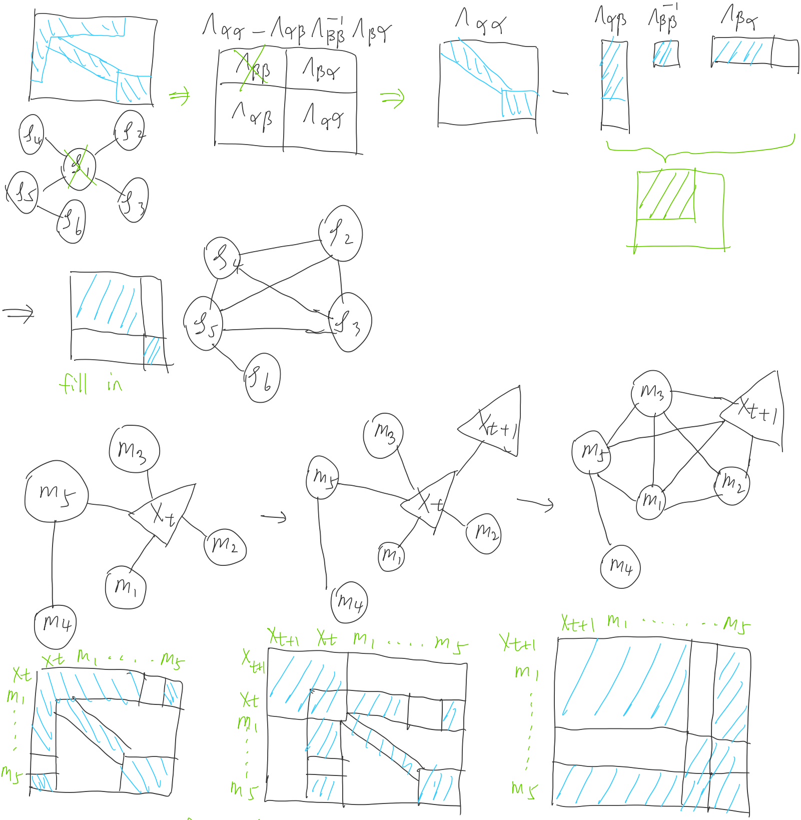
$$\Lambda_2 = J_2^T \Sigma_2^{-1} J = \begin{bmatrix} \left(\frac{\partial r_{13}}{\partial \xi_1}\right)^T \Sigma_2^{-1} \frac{\partial r_{13}}{\partial \xi_1} & 0 & \left(\frac{\partial r_{13}}{\partial \xi_1}\right)^T \Sigma_2^{-1} \frac{\partial r_{13}}{\partial \xi_3} & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ \left(\frac{\partial r_{13}}{\partial \xi_3}\right)^T \Sigma_2^{-1} \frac{\partial r_{13}}{\partial \xi_1} & 0 & \left(\frac{\partial r_{13}}{\partial \xi_3}\right)^T \Sigma_2^{-1} \frac{\partial r_{13}}{\partial \xi_3} & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

∴ Jacobian, info. matrix are sparse.



to remove ξ_1 in toy eg. using marginal: info. matrix change is below

$P(\xi_1, \dots \xi_6)$



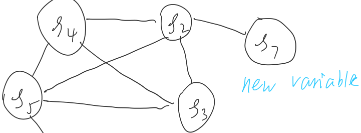
information filter (different from kalman filter)

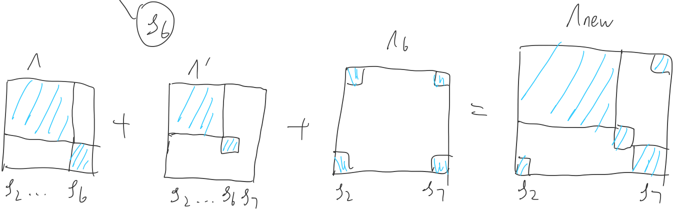
marginalization makes the information matrix dense, since the independent variables can become dependent.

FEJ

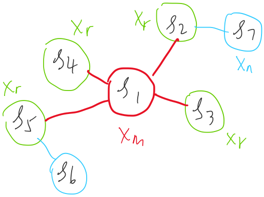
toy eg 3: remove ξ_1 and add new variable

new variable ξ_7 and ξ_2 has measurement and residual r_{27}.





review of toy eg 3:



marginalized variables

kept variables that are related with the marginalized variables

kept variables that are independent to the marginalized ones.

at $t \in [0, k]s$, variables are s_i, $i \in [1, 6]$.
at k', add new variable s_7.
at $t = k$, after LS, marginalize s_1, we call the marginalized variable X_m, the rest are X_i, $i \in [2, 5]$, denote it as X_r.
at $t = k'$, add s_7 (denote as X_n), begin new LS.

before marginalization, we have variable X_m and measurements S_m

$$b_m(k) = \begin{bmatrix} b_{mm}(k) \\ b_{mr}(k) \end{bmatrix} = -\sum_{(i,j) \in S_m} J_{ij}^T(k)\, \Sigma_{ij}^{-1}\, r_{ij}$$

$$\Lambda_m(k) = \begin{bmatrix} \Lambda_{mm}(k) & \Lambda_{mr}(k) \\ \Lambda_{rm}(k) & \Lambda_{rr}(k) \end{bmatrix} = \sum_{(i,j) \in S_m} J_{ij}^T(k)\, \Sigma_{ij}^{-1}\, J_{ij}(k)$$

after marginalization, measurements related to X_m passed to X_r:

$$b_p(k) = b_{mr}(k) - \Lambda_{rm}(k)\,\Lambda_{mm}^{-1}(k)\, b_{mm}(k)$$

$$\Lambda_p(k) = \Lambda_{rr}(k) - \Lambda_{rm}(k)\,\Lambda_{mm}^{-1}(k)\,\Lambda_{mr}(k)$$

p = prior, info. above is a prior for X_r.
we can derive a residual $r_p(k)$, jacobian $J_p(k)$ from $b_p(k)$, $\Lambda_p(k)$, $r_p(k)$, $b_p(k)$ will change, but $J_p(k)$ won't.

at $t = k'$, new residual r_{27}, and prior $b_p(k)$, $\Lambda_p(k)$ and residual r_{56} form a new LS:

$$b(k') = \Pi^T b_p(k) - \sum_{(i,j) \in S_a(k')} J_{ij}^T(k')\, \Sigma_{ij}^{-1}\, r_{ij}(k')$$

$$\Lambda(k') = \underbrace{\Pi^T \Lambda_p(k)\, \Pi}_{\text{prior}} + \underbrace{\sum_{(i,j)\in S_a(k')} J_{ij}^T(k')\, \Sigma_{ij}^{-1} J_{ij}(k')}_{}$$

where $\Pi = [I_{\dim x_r} \quad 0]$, S_a is measurements except those marginalized, eg r_{56}, r_{27}.

issues:

1. ∵ Jacobian in prior $\Lambda_p(k)$ cannot be updated
2. some variables in x_r are still related to other residuals, e.g. ξ_2, ξ_5.

the Jacobians of these residuals wrt ξ_2 will be computed at new ξ_2 values

Jacobians in prior are not changed.
Jacobians in current variables are updated.
} the linearization points to the same variable are not the same.

∴ nullspace of info. matrix is changed.

In BA:

$$\Lambda \delta x = b$$

$$\Lambda \delta x + N \delta x = b$$

change of δx in N will not change the residual ∴ there'll be multiple x satisfy the solution.

Eg. in DSO eg:

$$E = E_1 + E_2 = \underset{①}{(xy-1)^2} + \underset{②}{(xy-1)^2} \Rightarrow xy = 1.$$

if we obtain ① first, then ②:

$E_1' = E_1$ linearize around $(0.5, 1.4)$

$E_2' = E_2$, linearize around $(1.2, 0.5)$,

$E' = E_1' + E_2'$, nullspace disappears! $x\ y$ have a definite solution now.

Observability:

$z = h(\theta) + \varepsilon$, where

$z \in \mathbb{R}^n$, is measurement

$\theta \in \mathbb{R}^d$ is state variable

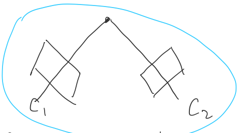
ε is observation noise

if the following holds, then θ is observable

$$\forall \theta, \forall \theta' \in \mathbb{R}^d, \{\theta \neq \theta'\} \Rightarrow \{h(\theta) \neq h(\theta')\}$$

for VO/SLAM, when we change the state, measurement does not change means cost function won't change, i.e the info. matrix Λ of LS has nullspace.

Mono-SLAM: 6 Dof + scale = 7 Dof unobservable



we move the system freely in 6Dof, i.e. times transformation matrix T, then the observations are not changed.

mono + imu. yaw + 3Dof = 4 Dof unobservable.

roll, pitch can be measured due to gravity,

scale can be measured due to accelerometer

first estimate jacobian:

Keep all the linearization points in all residuals the same, e.g. in toy eg 3, r_{27} to ξ_2 jcobian, linearization point of ξ_2 must be same with r_{12}. ie. use jacobian at k, instead of k'.