√SAM

→ ~~compared 2 EK~~

→ advntgs of √SAM

* in EKF SLAM, Σ or Σ^{-1} get change over tym.

in SAM, I is a remains sparse

I, A r compact reprs of map covar str.

→ most dramatic perf improvmnt is frm choosing a gud var ordering wen factorizing a mat.

QRD:

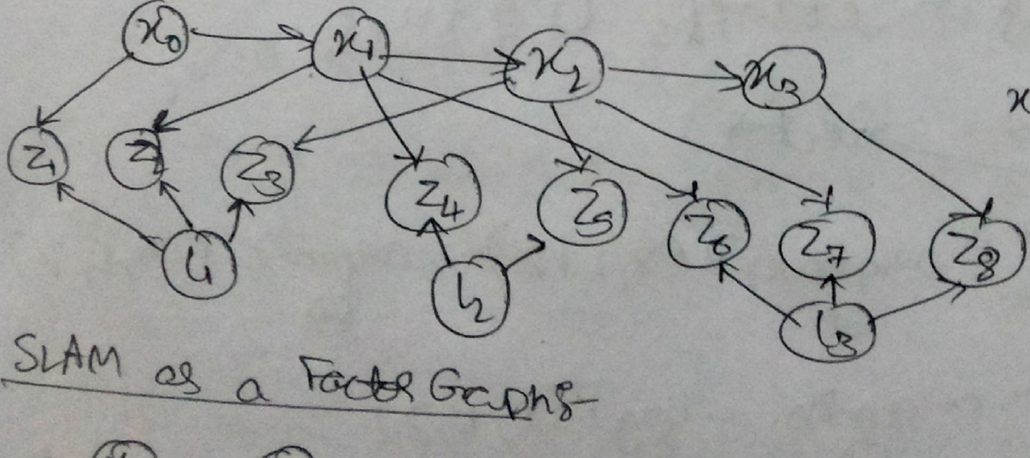
$A = QR$ Q → ortho R → upper Δlr

used to solve LLS.   QR Algo → an eig val algo.

SLAM as a Belief Net:

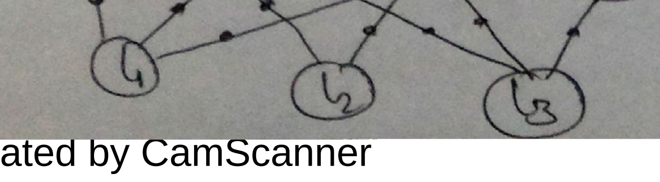
→ bel net → DAG encoding cndtnl independnc strctr of a set of vars, whr each var directly depnds only on its predecessors in d graph.

$$P(X, L, Z) = P(x_0) \underbrace{\prod_{i=1}^{M} P(x_i \mid x_{i-1}, u_i)}_{\text{motion model}} \underbrace{\prod_{k=1}^{K} P(z_k \mid x_{i_k}, l_{j_k})}_{\text{obs model}}$$



$$x_i = f_i(x_{i-1}, u_i) + w_i$$

SLAM as a Factor Graph:



$$P(\Theta) \propto \prod_i \phi_i(\theta_i) \prod_{\{i,j\}, i<j} \psi_{ij}(\theta_i, \theta_j)$$

Generated by CamScanner

$$P(\Theta) \propto \prod_i \phi_i(\theta_i) \prod_{\{i,j\}, i<j} \psi_{ij}(\theta_i, \theta_j)$$

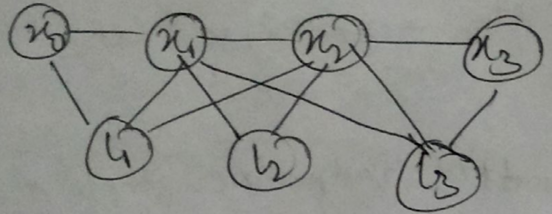
$\phi_i(\theta_i) \rightarrow$ encode a prior / single measuremt or constr at an unknown $\theta_i \in \Theta$, ~~whereas~~ +

$\psi_{ij}(\theta_i, \theta_j) \Rightarrow$ rel8 2 measremts or cnstrs dat involv d rltnshp b/w 2 unknowns θ_i, θ_j

$$\phi_0(x_0) \propto P(x_0) \qquad \psi_{(i-1)i}(x_{i-1}, x_i) \propto P(x_i | x_{i-1}, u_i)$$

$$\psi_{i_k j_k}(x_{i_k}, l_{j_k}) \propto P(z_k | x_{i_k}, l_{j_k})$$

SLAM as an MRF :-



Here, MRFs n Factor Graphs r equvt reprs.

SAM as an LS prblm :-

→ In SAM, we're concrnd with smoothing, rther dan filtering i.e., v wud lyk 2 recover d MAP estimt of d entire traj $X \triangleq \{x_i\}$ n d map $L \triangleq \{l_j\}$ given $Z \triangleq \{z_k\}$ n ctrl i/ps $U \triangleq \{u_i\}$.

Say $\underbrace{\Theta \triangleq (X, L)}_{\text{vector}}$

MAP estim8 $\quad \Theta^* \triangleq \underset{\Theta}{\operatorname{argmax}} P(X, L | Z) = \underset{\Theta}{\operatorname{argmax}} P(X, L, Z)$

$$= \underset{\Theta}{\operatorname{argmin}} -\log P(X, L, Z)$$

$\Theta^* \triangleq \underset{\Theta}{\operatorname{argmin}} \left\{ \sum_{i=1} \right.$

$x_i = f_i(x_{i-1}, u_i) + w_i \Leftrightarrow \cancel{P(x_i|x_{i-1}, u_i) \propto \exp\{-\frac{1}{2}\|}$

$\Rightarrow P(x_i|x_{i-1}, u_i) \propto \exp\left\{-\frac{1}{2}\| f_i(x_{i-1}, u_i) - x_i\|^2_{\Lambda_i}\right\}$

$\Lambda_i \rightarrow$ prcss noise covar mat $\qquad w_i \sim N(0, \Lambda_i)$

$z_k = h(x_{i_k}, l_{j_k}) + v_k \Leftrightarrow P(z_k|x_{i_k}, l_{j_k}) \propto \exp\left\{-\frac{1}{2}\|h_k(x_{i_k}, l_{j_k}) - z_k\|^2_{\Sigma_k}\right\}$

$\Sigma_k \rightarrow$ obs covar mat $\qquad v_k \sim N(0, \Sigma_k)$

$\|e\|^2_\Sigma = e^T \Sigma^{-1} e \rightarrow$ squared Mahalanobis dist

$\Sigma \rightarrow$ covar mat $\Rightarrow$ PSD $\Rightarrow \Sigma^{-1}$ xists.

MAP estimn:

$$\Theta^* \triangleq \underset{\Theta}{\arg\max}\, P(X, L|Z) = \underset{\Theta}{\arg\max}\, P(X, L, Z)$$

$$= \underset{\Theta}{\arg\min}\,(-\log P(X, L, Z))$$

$$\Theta^* \triangleq \arg\min \left\{ \sum_{i=1}^{M} \|f_i(x_{i-1}, u_i) - x_i\|^2_{\Lambda_i} + \sum_{k=1}^{K} \|h_k(x_{i_k}, l_{j_k}) - z_k\|^2_{\Sigma_k} \right\}$$

(v assume dat d prior $P(x_0)$ is given, n hence const)

→ If d prcss models f_i n measurement eqns h_k r nonlnr n a gud linearizatn pt isn't avlbl, non lin optn methods such as Gauss-Newton itrtns or d Levenberg-Marquardt algo 'll solve a succession of lnr approxm^ns 2 d above eqn.

→ dis is smlr 2 d EKF approach 2 SLAM, bt allows for itrtng mltpl tyms 2 cnvrgnc while cntrlling in which rgn 1 is willng 2 trust d linear assumption (hence dese methods r often calld rgn-trust methods).

$f_i(x_{i-1}, u_i) - x_i \approx \{f_i(x_{i-1}^0, u_i) - x_i^0\} - \{F_i^{i-1}\delta x_{i-1} - \delta x_i\}$

$\approx \{f_i(x_{i-1}^0, u_i) + F_i^{i-1}\delta x_{i-1}\} - \{x_i^0 + \delta x_i\}$

$= \{F_i^{i-1}\delta x_{i-1} - \delta x_i\} - a_i$

$F_i^{i-1} \rightarrow$ Jacob of $f_i(.)$ at d lin^{zn} pt x_{i-1}^0

dfned by $\quad F_i^{i-1} \triangleq \left.\dfrac{\partial f_i(x_{i-1}, u_i)}{\partial x_{i-1}}\right|_{x_{i-1}^0}$

$a_i \triangleq -\left(f_i(x_{i-1}^0, u_i) - x_i^0\right) \rightarrow$ odom pred^n err

$h_k(x_{i_k}, l_{j_k}) - z_k \approx \{h_k(x_{i_k}^0, l_{j_k}^0) + H_k^{i_k}\delta x_{i_k} + J_k^{i_k}\delta l_{j_k}\} - z_k$

$= \{H_k^{i_k}\delta x_{i_k} + J_k^{i_k}\delta l_{j_k}\} - c_k \rightarrow$ measmnt pred^n err

So, $\delta^* = \underset{\delta}{\operatorname{argmin}} \left\{ \sum_{i=1}^{M} \|F_i^{i-1}\delta x_{i-1} + G_i^i \delta x_i - a_i\|_{\Lambda_i}^2 + \sum_{k=1}^{K} \|H_k^{i_k}\delta x_{i_k} + J_k^{j_k}\delta l_{j_k} - c_k\|_{\Sigma_k}^2 \right\}$

$G_i^i = -I_{d\times d} \qquad x_i \in \mathbb{R}^d$

$\|e\|_\Sigma^2 \triangleq e^T\Sigma^{-1}e = e^T\Sigma^{-\frac{1}{2}}\Sigma^{-\frac{1}{2}}e = e^T\Sigma^{-\frac{T}{2}}\Sigma^{-\frac{1}{2}}e$

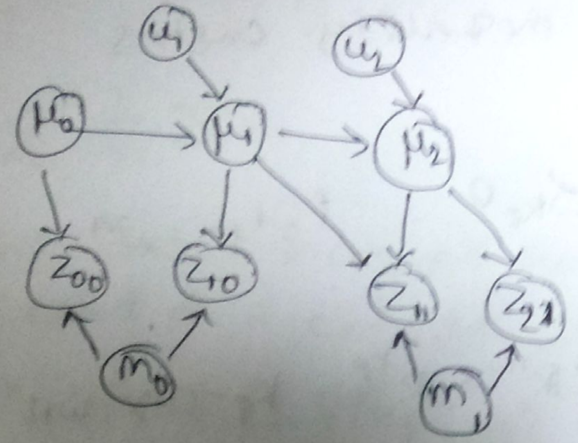
$= (\Sigma^{-\frac{1}{2}}e)^T(\Sigma^{-\frac{1}{2}}e) = \|\Sigma^{-T/2}e\|_2^2$

Generated by CamScanner

26/10/15
MON

MR | Lec#16

numerator-layout vs denominator-layout



Consider a 3 bot-step 2-landmark situation

v derive SAM eqns f dem.

~~why do v consider only u_1 n not u_2 ?~~

z_{ij} — i → time instant, j → landmark #

find $\mu^*, \mu_m^* = \underset{\mu, \mu_m}{\arg\max}\ P(\mu, \mu_m \mid u, z)$

$= \underset{\mu, \mu_m}{\arg\max}\ P(\mu, \mu_m, u, z)$

$\cancel{\propto P(\mu_2 \mid \mu_1, u_1)\ P(z_{21}, z_{11} \mid m_1, \mu_2)}$
$\cancel{P(\mu_1 \mid \mu_0, u_0)\ P(z_{11} \mid m_1)\ P(z_0 \mid m_0)}$
$\cancel{P(\mu_0)\ P(z_{00} \mid m_0) \ldots P(z_{10} \mid m_0)}$

$\propto P(\mu_2 \mid \mu_1, u_2)\ P(z_{21} \mid \mu_2, m_1)\ P(z_{11} \mid \mu_2, m_1)$
$P(\mu_1 \mid \mu_0, u_1)\ P(z_{11} \mid \mu_1, m_1)\ P(z_{10} \mid \mu_1, m_0)$
$P(\mu_0)\ P(z_{00} \mid \mu_0, m_0)$

assumed 2 b known beforehand

$\underset{\mu_0, \mu_1, \mu_2, m_0, m_1}{\arg\max}\ P(\mu_2, \mu_1, \mu_0, z_{00}, z_{10}, z_{11}, z_{21}, m_0, m_1, u_1, u_2)$

Generated by CamScanner

cvx! $\boxed{\text{Lec\#07}}$

Gnrlized Inequality Cnstrnts:-

→ cvx prblm with gnrlized inequality cnstrs

$$\min \; f_0(x)$$
$$\text{subj to} \quad f_i(x) \preceq_{K_i} 0 \qquad i = 1, \ldots, m$$
$$Ax = b$$

$f_0 : \mathbb{R}^n \to \mathbb{R}$ cvx; $\quad f_i : \mathbb{R}^n \to \mathbb{R}^{k_i}$ K_i-cvx wrt proper cone K_i

→ same prprties as std cvx prblm

→ conic form prblm (spl case with affine obj n cnstrs

$$\min \; c^T x$$
$$\text{subj to} \quad Fx + g \preceq_K 0$$
$$Ax = b$$

extnds linear progg ($K = \mathbb{R}^m_+$) to non polyhedral cones

---

$x_1, x_2 \qquad P(x_1, x_2) = P(x_2 | x_1)\, P(x_1)$

IID ⇒ $P(x_2 | x_1) = P(x_2)$

$$P(x_1, x_2, \ldots, x_n ; \mu) = \prod_{i=1}^{n} P(x_i ; \mu)$$

$(0.3)^2 (0.7)^3 \;\to$ if $P(H) = 0.3$

ML → choose μ dat maximizes d likelihood fn

$$P(X ; \mu) = \prod_{i=1}^{n} P(x_i ; \mu) = \prod_{i=1}^{n} \mu^{x_i} (1-\mu)^{1-x_i}$$

$$\log P(X ; \mu) = \sum x_i \log \mu + (1 - x_i) \log (1-\mu)]$$

$\frac{\partial}{\partial \mu} \log p(X;\mu) = \sum \left(\frac{x_i}{\mu} - \frac{1-x_i}{1-\mu} \right) = 0 \Rightarrow \frac{1}{\mu} \sum x_i = \frac{1}{1-\mu} \sum (1-x_i)$

$\Rightarrow \frac{1-\mu}{\mu} = \frac{\sum 1-x}{\sum x} \Rightarrow \frac{1}{\mu} - 1 = \frac{n}{\sum x} - 1$

$\Rightarrow \hat{\mu}_{ML} = \frac{1}{n} \sum x_i$

ML estimator doesn't incorporate any prior knowledge n doesn't generate an estimate of the certainty of its results.

$$p(X;\mu) = \prod_i p(x_i;\mu) = \prod_i \mu^{x_i} (1-\mu)^{(1-x_i)}$$

$$\log p(X;\mu) = \sum_i x_i \log \mu + (1-x_i) \log (1-\mu)$$

$$\frac{\partial \log p(X;\mu)}{\partial \mu} = \sum_i \left(\frac{x_i}{\mu} - \frac{1-x_i}{1-\mu} \right) = 0$$

$$\Rightarrow \frac{1}{\mu} \sum_i x_i = \frac{1}{1-\mu} \sum_i (1-x_i)$$

$$\Rightarrow \frac{1-\mu}{\mu} = \frac{\sum_i (1-x_i)}{\sum_i x_i} \Rightarrow \frac{1}{\mu} - 1 = \frac{\sum_i 1}{\sum_i x_i} - 1$$

$$\Rightarrow \hat{\mu}_{ML} = \frac{\sum_i x_i}{n}$$

HTHT $\Rightarrow \hat{\mu} = 0.5$

HTTT $\Rightarrow \hat{\mu} = 0.25$

TTTT $\Rightarrow \hat{\mu} = 0.0$

MAP:

$$p(\mu|X) = \frac{p(X|\mu)\, p(\mu)}{p(X)} \qquad \hat{\mu}_{MAP} = \underset{\mu}{\operatorname{argmax}}\; p(\mu|X)$$

$p(X)$ doesn't depend on μ.

So $\hat{\mu}_{MAP} = \underset{\mu}{\operatorname{argmax}}\; p(\mu|X) = \underset{\mu}{\operatorname{argmax}} \frac{p(X|\mu)\, p(\mu)}{p(X)}$

$= \underset{\mu}{\operatorname{argmax}}\; p(X|\mu)\, p(\mu)$

Generated by CamScanner

Beta dist'n

→ appropri8 2 xprss prior bel abt a Bernoulli dist.

$$P(\mu) = \frac{1}{B(\alpha, \beta)} \mu^{\alpha-1}(1-\mu)^{\beta-1}$$

$$B(\alpha, \beta) = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)}$$

$$\hat{\mu}_{MAP} = \arg\max_{\mu} P(X|\mu)\, P(\mu) = \arg\max_{\mu} \prod_i P(x_i|\mu)\, P(\mu)$$

$$= \arg\max_{\mu} \sum_i \{\log P(x_i|\mu) + \log P(\mu)\}$$

$$\underbrace{P(\mu|X)}_{\text{posterior}} \propto \underbrace{P(X|\mu)}_{\text{likelihood}}\ \underbrace{P(\mu)}_{\text{prior}}$$

ML Assg ~~ttor~~

MNIST → 60k train / 10k test

digits (size-normalized n centrd

DL $\boxed{\text{Lec\#04}}$

Nonlnr RidgeRegr : Risk, Regularizn, n Cross-Validn

~~CNN~~

~~Linear CNN~~

$n \times d$ ~~data~~

$X \to n \times d$

$$\hat{\theta} = \underbrace{\underbrace{(X^T X)^{-1}}_{d \times d}\ \underbrace{X^T}_{d \times n}\ \underbrace{y}_{n \times 1}}_{d \times d}$$

$$\hat{\theta} = (X^T X + \delta^2 I_d)^{-1} X^T y$$

$$J(\theta) = (y - X\theta)^T (y - X\theta) + \delta^2 \theta^T \theta$$

$$\frac{\partial}{\partial \theta} J(\theta) = \frac{\partial}{\partial \theta} \left\{ \theta^T X^T X \theta - 2y^T X \theta + y^T y + \delta^2 \theta^T \theta \right\}$$

Generated by CamScanner

$= 2X^TX\theta - 2y^TX + 2\delta^2 I_d \theta$

$= 2(X^TX + \delta^2 I_d)\theta - 2y^TX = 0$

$\hat{\theta}_{ridge} = (X^TX + \delta^2 I_d)^{-1} X^Ty$

$\hat{\theta}_{LS} = (X^TX)^{-1} X^Ty$

$X^TX\hat{\theta}_{ridge} + \delta^2 I_d \hat{\theta}_{ridge} = X^Ty$

$\Rightarrow \hat{\theta}_{ridge} = \hat{\theta}_{LS} - (X^TX)^{-1}\delta^2 I_d \hat{\theta}_{ridge}$

$= \hat{\theta}_{LS} - \delta^2 (X^TX)^{-1} I_d \cancel{(X^TX + \delta^2 I)^{-1}} \hat{\theta}_{ridge}$

$J(\theta) = (y - X\theta)^T (y - X\theta) + \delta^2 \theta^T\theta$

Kernel Regr → u r thing d X_s' y not d y's?

$e^{-\frac{1}{2}}$ —— —— $P_c(\bar{c}|S)$ c → crrspondnc
S → gnd plane

Sampling prior → $P_c(c|S)$

$\bar{x}$ → homo repr of 2D pt in img coords

$\vec{x}$ → crrsponding 3D pt. $\vec{x}' = R\vec{x} + \vec{S}$

$\vec{n}^T\vec{x} = d$

$\bar{x}' = H\bar{x} = K(R + \vec{t}\cdot\vec{n}^T)K^{-1}\bar{x}$

$x = KX$ $n^TX = d$ x'

$X' = RX + s$ $n \cdot X = d$

$\bar{t} = \frac{\bar{S}}{d}$ $x' = Hx = K(R + n \cdot t)K^{-1}x$

Generated by CamScanner

Smoothing ⇒ recover d MAP estimate for a entire traj.

$X \triangleq \{x_i\}$ n map $L \triangleq \{l_j\}$, given measurements

$Z \triangleq \{z_k\}$ n ctrl i/p s $U \triangleq \{u_i\}$

$\Theta \triangleq (X, L)$

MAP estimate $\Theta^* \triangleq \underset{\Theta}{\operatorname{argmax}}\, P(X, L | Z) = \underset{\Theta}{\operatorname{argmax}}\, P(X, L, Z)$

$$= \underset{\Theta}{\operatorname{argmin}}\, -\log P(X, L, Z)$$

$$x_i = f_i(x_{i-1}, u_i) + w_i \Leftrightarrow P(x_i | x_{i-1}, u_i) \propto \exp\left\{-\frac{1}{2}\|f_i(x_{i-1}, u_i) - x_i\|^2_{\Lambda_i}\right\}$$

$$z_k = h_k(x_{i_k}, l_{j_k}) + v_k \Leftrightarrow P(z_k | x_{i_k}, l_{j_k}) \propto \exp\left\{-\frac{1}{2}\|h_k(x_{i_k}, l_{j_k}) - z_k\|^2_{\Sigma_k}\right\}$$

$$\Theta^* \triangleq \underset{\Theta}{\operatorname{argmin}} \left\{ \sum_{i=1}^{M} \|f_i(x_{i-1}, u_i) - x_i\|^2_{\Lambda_i} + \sum_{j=1}^{N} \|h_k(x_{i_k}, l_{j_k}) - z_k\|^2_{\Sigma_k} \right\}$$

$$f_i(x_{i-1}, u_i) - x_i \approx f_i(x^0_{i-1}, u_i) - x^0_i + F_i^{i-1}\delta x_{i-1} - \delta x_i$$

$$= \left(F_i^{i-1}\delta x_{i-1} - \delta x_i\right) - a_i$$

$$h_k(x_{i_k}, l_{j_k}) - z_k \approx h_k(x^0_{i_k}, l^0_{j_k}) - z^0_k + H_k^{i_k}\delta x_{i_k} + J_k^{j_k}\delta l_{j_k}$$

$$= \left\{H_k^{i_k}\delta x_{i_k} + J_k^{j_k}\delta l_{j_k}\right\} - c_k$$

$$\delta^* = \underset{\delta}{\operatorname{argmin}} \left\{ \sum_{i=1}^{M} \|F_i^{i-1}\delta x_{i-1} + G_i^i \delta x_i - a_i\|^2_{\Lambda_i} + \sum_{j=1}^{N} \|H_k^{i_k}\delta x_{i_k} + J_k^{j_k}\delta l_{j_k} - c_k\|^2_{\Sigma_k} \right\}$$

$$\|e\|^2_{\Sigma} \triangleq e^T \Sigma^{-1} e = e^T \Sigma^{-\frac{T}{2}} \Sigma^{-\frac{1}{2}} e = \|\Sigma^{-T/2} e\|_2^2$$

$$\delta^* = \underset{\delta}{\operatorname{argmin}}\, \|A\delta - b\|_2^2$$

Generated by CamScanner

$$x_{t+1} = A x_t + B u_{t+1} + \cup \nearrow Q^R = f(x_t, u_{t+1})$$

$$x_{t+1} \approx x^0_{t+1} + F \cdot - + G \cdot -$$

$$\Sigma' = F \Sigma^{-1} F^T + G \Sigma_u G^T$$

$$F (\Sigma')^{-1} F^T$$

$$FF^T + GG^T$$

$$F \Sigma^{-1/2} \Sigma^{-1/2} F$$

$$= (\Sigma^{-1/2} F)^T (\Sigma^{-1/2} F) = \| \Sigma^{-1/2} F \|^2 \cdot \quad F^T F$$

$$J^T J$$

$$\delta^* = \arg\min_\delta \left\{ \sum_{i=1}^{M} \| (F_p^{i-1} \delta x_{i-1} + G_p^i \delta x_p - a_i) \|^2 \right.$$

$$\left. + \sum_{k=1}^{K} \| H_k^{p_k} \delta x_{p_k} + J_k^{j_k} \delta l_{j_k} - c_k \|^2_{\Sigma_k} \right\}$$

$M = 3, \quad N = 2, \quad K = 4$

$$\delta^* = \arg\min_\delta \left\{ \left[(G_1^1 \delta x_1 - a_1) + (F_2^1 \delta x_1 + G_2^2 \delta x_2 - a_2) \right. \right.$$

$$\left. + (F_3^2 \delta x_2 + G_3^3 \delta x_3 - a_3) \right] + \left[(H_1^1 \delta x_1 + J_1^1 \delta l_1 - c_1) \right.$$

$$\left. \left. + (H_2^1 + \quad - \cdots) + - - - \right] \right\}$$

$$\| A\delta - b \|_2^2 \qquad \delta = [\delta x_1 \;\; \delta x_2 \;\; \delta x_3 \;\; \delta l_1 \;\; \delta l_2]^T$$

$$(N d_x + K d_2) \times (N d_x + M d_l)$$

$$(A\delta - b)^T (A\delta - b) = \delta^T A^T A \delta - 2 \delta^T A^T b + b^T b$$

$$A^T A \delta^* = A^T b \qquad \mathcal{I} \triangleq A^T A = R^T R \qquad R^T R \delta^* = A^T b$$

$$R^T y = A^T b, \quad R\delta^* = y$$

$$Q^T A = \begin{bmatrix} R \\ 0 \end{bmatrix} \qquad Q^T b = \begin{bmatrix} d \\ e \end{bmatrix} \qquad \mathcal{I} = \begin{bmatrix} A_x^T A_x & \mathcal{I}_{xL} \\ \mathcal{I}_{xL}^T & A_L^T A_L \end{bmatrix}$$

Generated by CamScanner

Batch $\sqrt{SAM}$

→ Build A n b (A → measurement Jacob, b → RHS)

→ Find a gud col ordering P, n reorder $A_p \xleftarrow{P} A$

→ Solve $\delta_p^* = \arg\min_{\delta} \|A_p\delta_p - b\|_2^2$ using either d Cholesky or QR factorizⁿ method

→ Recover d optimal solⁿ by $\delta \xleftarrow{r} \delta_p$ $(r = p^{-1})$

$$\|A\delta - b\|_2^2 = \|Q^TA\delta - Q^Tb\|_2^2 = \|R\delta - d\|_2^2 + \|e\|_2^2$$

$Q^TQ = I$

$R\delta = d$ → solve via back substitution.

$A = QR \Rightarrow Q^TA = R$

$\det(A)$

$$\begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$$

$$\delta^TA^TA\delta - 2\delta^TA^Tb + b^Tb$$

$a^Tb = b^Ta$

$$\delta^TA^TQQ^TA\delta - 2b^TQQ^TA\delta + b^TQQ^Tb$$

$\cancel{-2b^TQ}$

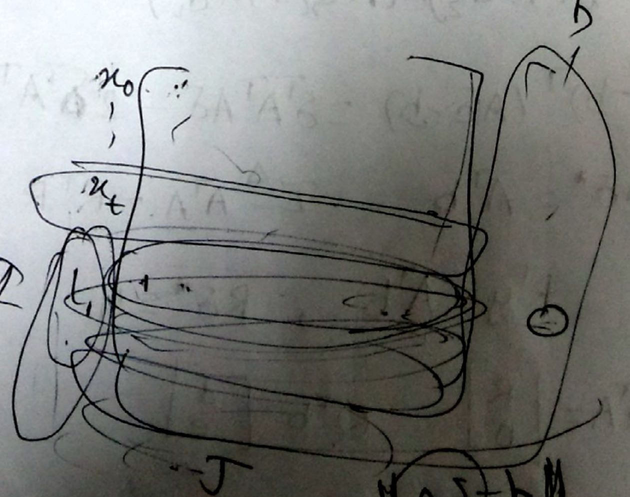
$$\delta^TA^Tb = b^TA\delta$$

$\cancel{(Q^TA\delta)}$

$$(Q^Tx)^T(Q^Tx)$$

$$\|Q^T(A\delta - b)\|_2^2$$

$$\|Q^T\|^2 = (Q^T)^TQ^T = I$$



$\|A\delta - b\|_2^2$

$J^TJ \rightarrow A$

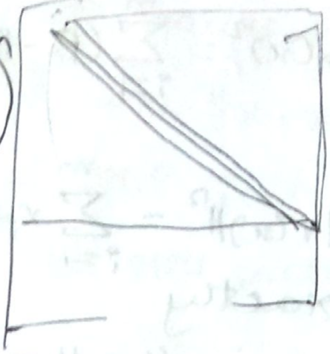
Generated by CamScanner

$$\|Q^T(A\delta - b)\|^2 = \|Q^T\|^2 \; \|(A\delta - b)\|$$

$$Q^T Q$$

$A = Q R$

$m \times n \quad m \times m \quad m \times n$



$\begin{bmatrix} R \\ 0 \end{bmatrix} \to m \times n$

$n \times n$

$m \times n$

$\phi(x; t)$

$\cancel{m \times n}\ m > n, \; r_j(x) = \phi(x; t_j) - y_j$

GN, LM

$$\nabla f(x_1, x_2) = \left(\frac{\partial f}{\partial x_1}, \frac{\partial f}{\partial x_2}\right)$$

$$H(f(x_1, x_2)) = \begin{matrix} \frac{\partial^2 f}{\partial x_1^2} & \frac{\partial f}{\partial x_1 \partial x_2} \\ \frac{\partial f}{\partial x_2 \partial x_1} & \frac{\partial f}{\partial x_2^2} \end{matrix}$$

$$[f + G]^T \lambda_i^{-1} [f + G]$$

$$\cancel{f \lambda_1} \left[f^T \lambda_i^{-1} + G^T \lambda_i^{-1}\right] \cdot [f + G].$$

$$= \cancel{(\lambda_1)^T}$$

$$f(x_0 + \delta) = f(x_0) + f'(x_0)\,\delta + \frac{1}{2} f''(x_0)\,\delta^2 + \ldots$$

$$f(x_0 + \delta) \approx f(x_0) + f'(x_0)\,\delta \qquad x_1 = x_0 + \delta = x_0 - \frac{f(x_0)}{f'(x_0)}$$

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)} \qquad x_{n+1} = x_n - \frac{f'(x_n)}{f''(x_n)}$$

$$\nabla f(x) = J(x)^T r(x)$$

$$\nabla f(x) = \sum_{j=1}^{m} r_j(x)\, \nabla r_j(x) = J(x)^T r(x)$$

$$f(x) = \frac{1}{2} \sum_{j=1}^{m} r_j^2(x)$$

$$\nabla f(x) = \sum_{j=1}^{m} r_j(x)\, \nabla r_j(x) = J(x)^T r(x)$$

$$\nabla^2 f(x) = \sum_{j=1}^{m} (\nabla r_j(x))^2 + \sum_{j=1}^{m} r_j(x)\, \nabla^2 r_j(x)$$

$$= J(x)^T J(x) + \sum_j r_j(x)\, \nabla^2 r_j(x)$$

Generated by CamScanner