Computationally Efficient Belief Space Planning via Augmented Matrix Determinant Lemma and Re-Use of Calculations
Supplementary Material

Technical Report ANPL-2017-01

Dmitry Kopitkov and Vadim Indelman

This document provides supplementary material to the paper [1]. Therefore, it should not be considered a self-contained document, but instead regarded as an appendix of [1], and cited as:

"D. Kopitkov and V. Indelman, *Computationally Efficient Belief Space Planning via Augmented Matrix Determinant Lemma and Re-Use of Calculations*, (Supplementary Material, ANPL-2017-01), IEEE Robotics and Automation Letters (RA-L), accepted."

Throughout this report, standard notations are used to refer to equations from [1] (e.g. Eq. (5)), while equations introduced herein are represented by the corresponding Appendix letter, e.g. Eqs. (A1) and (A2).

This document is organized as follows: Appendices A, B and C provide proofs for Lemmas 1, 2 and 3 respectively.

Appendix A: Proof of Lemma 1

Problem definition: Given a positive definite and symmetric matrix $\Lambda \in \mathbb{R}^{n\times n}$ (e.g. a prior information matrix) and its inverse Σ (prior covariance matrix), first Λ is augmented by k zero rows and columns and the result is stored in Λ^{Aug}. Then we have matrix $A \in \mathbb{R}^{m\times(n+k)}$ and calculate $\Lambda^{+} = \Lambda^{Aug} + A^T \cdot A$ (see Figure 1). We would like to express the determinant of Λ^{+} in terms of Λ and Σ.

D. Kopitkov is with the Technion Autonomous Systems Program (TASP), Technion - Israel Institute of Technology, Haifa 32000, Israel, dimkak@tx.technion.ac.il . V. Indelman is with the Department of Aerospace Engineering, Technion - Israel Institute of Technology, Haifa 32000, Israel, vadim.indelman@technion.ac.il. This work was supported by the Israel Science Foundation.

We start by modeling the matrix Λ^{Aug} through Σ. By introducing k new variables, before adding any new constraints involving these variables, we can say that new variables are uncorrelated with old variables, and their uncertainty is infinite (nothing yet is known about them). Then the appropriate covariance matrix after augmentation, Σ^{Aug}, can just be created by adding k zero rows and columns to Σ, and setting new diagonal entries with parameter θ, noting that $\theta \to \infty$:

$$\Sigma^{Aug} = \begin{bmatrix} \Sigma & 0 \\ 0 & \theta \cdot I \end{bmatrix}. \tag{A1}$$

Next, note that inverse of Σ^{Aug} is given by the following expression:

$$(\Sigma^{Aug})^{-1} = \begin{bmatrix} \Lambda & 0 \\ 0 & \epsilon \cdot I \end{bmatrix}. \tag{A2}$$

where $\epsilon \doteq \frac{1}{\theta}$. Taking limit $\epsilon \to 0$ into account, we can see that the above equation converges to Λ^{Aug} as it was defined above. Then in the limit we will have that $(\Lambda^{Aug})^{-1} = \Sigma^{Aug}$. Also note that $\epsilon \to 0$, even that it never becomes zero, $\epsilon \neq 0$, thus if needed we can divide by ϵ without worry.

Taking into account the limit of ϵ, expressing Λ^{Aug} through Eq. (A2) will not change the problem definition. But such a model allows to inverse Λ^{Aug}:

$$(\Lambda^{Aug})^{-1} = \Sigma^{Aug} = \begin{bmatrix} \Sigma & 0 \\ 0 & \theta \cdot I \end{bmatrix}, \tag{A3}$$

and therefore to use the generalized matrix determinant lemma [2]:

$$\begin{aligned} \left|\Lambda^{+}\right| = \left|\Lambda^{Aug}\right| \cdot \left|I_m + A \cdot \Sigma^{Aug} \cdot A^T\right| = \\ \left|\Lambda\right| \cdot \epsilon^k \cdot \left|I_m + C \cdot \Sigma \cdot C^T + \theta \cdot D \cdot D^T\right| \end{aligned} \tag{A4}$$

where matrices $C \in \mathbb{R}^{m \times n}$ and $D \in \mathbb{R}^{m \times k}$ are constructed from A by retrieving columns of only old n variables and of only new k variables, respectively (see Figure 2a).

Using the matrix determinant lemma once more, we get:

$$\left|\Lambda^{+}\right| = \left|\Lambda\right| \cdot \epsilon^k \cdot \left|R\right| \cdot \left|I_k + \theta \cdot D^T \cdot R^{-1} \cdot D\right| \tag{A5}$$

where $R \doteq I_m + C \cdot \Sigma \cdot C^T$.

Moving ϵ inside the last determinant term, we have:

$$\left|\Lambda^{+}\right| = \left|\Lambda\right| \cdot \left|R\right| \cdot \left|\epsilon \cdot I_k + \epsilon \cdot \theta \cdot D^T \cdot R^{-1} \cdot D\right| \tag{A6}$$

Recalling that $\epsilon \to 0$ and $\epsilon \cdot \theta = 1$, we will get to:

$$\left|\Lambda^{+}\right| = \left|\Lambda\right| \cdot \left|R\right| \cdot \left|D^T \cdot R^{-1} \cdot D\right| \tag{A7}$$

And the augmented determinant ratio will be:

$$\begin{aligned} \frac{\left|\Lambda^{+}\right|}{\left|\Lambda\right|} = \left|I_m + C \cdot \Sigma \cdot C^T\right| \cdot \left|D^T \cdot (I_m + C \cdot \Sigma \cdot C^T)^{-1} \cdot D\right| = \\ \left|R\right| \cdot \left|D^T \cdot R^{-1} \cdot D\right| \end{aligned} \tag{A8}$$

$\blacksquare$

Appendix B: Proof of Lemma 2

Consider the scenario of focused BSP where the focused set X^F_{k+L} contains only newly added variables as defined in Section III-C1, with appropriate illustration shown in Figure 2b.

First let us overview the various partitions of Jacobian A which are relevant to our current problem (Figure 2b). C, D, C^I and $C^{\neg I}$ were already introduced in [1]. Further, we can partition D into D^F - columns of new variables that are focused $Y^F \equiv X^F_{k+L} \in \mathbb{R}^{n_F}$, and D^U - columns of new unfocused variables Y^U. Considering the figure, the set of all unfocused variables in X_{k+L} will be $X^R_{k+L} \doteq \{X \cup Y^U\} \in \mathbb{R}^{n_R}$, such that $N = n_F + n_R$, providing another A's partition $A_R = [C, D^U]$.

Next, we partition the posterior information matrix Λ_{k+L} respectively to the defined above sets X^F_{k+L} and X^R_{k+L} as

$$\Lambda_{k+L} = \begin{bmatrix} \Lambda^R_{k+L} & \Lambda^{R/F}_{k+L} \\ (\Lambda^{R/F}_{k+L})^T & \Lambda^F_{k+L} \end{bmatrix}. \tag{B9}$$

As was shown in previous paper [3], determinant of the marginal covariance of X^F_{k+L} can be calculated through:

$$\left|\Sigma^F_{k+L}\right| = \frac{\left|\Lambda^R_{k+L}\right|}{\left|\Lambda_{k+L}\right|}. \tag{B10}$$

Now let us focus on Λ^R_{k+L} term from the right side. From Eq. (5) we can see that partition of posterior information matrix Λ^R_{k+L} can be calculated as:

$$\Lambda^R_{k+L} = \Lambda^{Aug,R}_k + A^T_R A_R \tag{B11}$$

where $\Lambda^{Aug,R}_k$ can be constructed by augmenting Λ_k with zero rows and columns in number of Y^U's dimension (see Figure 2b). The above equation has augmented determinant form as defined in Section III-A, and so the augmented determinant lemma can be applied on it. Using Eq. (10) we have:

$$\frac{\left|\Lambda^R_{k+L}\right|}{\left|\Lambda_k\right|} = \left|P\right| \cdot \left|(D^U)^T \cdot P^{-1} \cdot D^U\right| \tag{B12}$$

where P is defined in Eq. (13).

Next, dividing Eq. (B12) by Eq. (11), we get

$$\left|\Sigma^F_{k+L}\right| = \frac{\left|\Lambda^R_{k+L}\right|}{\left|\Lambda_{k+L}\right|} = \frac{\left|(D^U)^T \cdot P^{-1} \cdot D^U\right|}{\left|D^T \cdot P^{-1} \cdot D\right|}, \tag{B13}$$

and posterior entropy of X^F_{k+L} is given by

$$J^F_{\mathcal{H}}(a) = \frac{n_F}{2} \cdot (1 + \ln(2\pi)) + \frac{1}{2} \ln \left|(D^U)^T \cdot P^{-1} \cdot D^U\right| - \frac{1}{2} \ln \left|D^T \cdot P^{-1} \cdot D\right|. \tag{B14}$$

$\blacksquare$

Note that the variables inside information matrices do not have to be ordered in any particular way, and that the provided above proof is correct for any ordering whatsoever.

Appendix C: Proof of Lemma 3

Consider the scenario of focused BSP where the focused set X_{k+L}^F contains only old variables, with appropriate illustration shown in Figure 2c and with various partitions of Jacobian A defined in Section III-C2.

First, let us look again over relevant partitions of Jacobian A (Figure 2c). The C, D, C^I and $C^{\neg I}$ were already introduced inside the paper [1]. From the figure we can see that $C^{\neg I}$ can further be separated into $C^{\neg I,U}$ - columns of old variables that are both not involved and unfocused ($X^{\neg I,U}$), and $C^{\neg I,F}$ - columns of old variables that are both not involved and focused ($X^{\neg I,F}$). Additionally, C^I can be partitioned into $C^{I,U}$ - columns of old variables that are both involved and unfocused ($X^{I,U}$), and $C^{I,F}$ - columns of old variables that are both involved and focused ($X^{I,F}$) (see Table I). The set of focused variables is then $X_{k+L}^F = \{X^{\neg I,F} \cup X^{I,F}\} \in \mathbb{R}^{n_F}$, containing both involved and not involved variables. We will denote $X_{k+L}^F \doteq X_k^F$ to remind us that focused set of variables is part of both X_{k+L} and X_k.

Likewise, the set of all remained, unfocused variables is $X_{k+L}^R \doteq \{X^{\neg I,U} \cup X^{I,U} \cup Y\} \in \mathbb{R}^{n_R}$, containing all new variables and some of old ones (which can be involved or not involved), and providing A's partition $A_R = [C^{\neg I,U}, C^{I,U}, D]$. Moreover, for purpose of simplification of coming equations we'll denote set of old variables inside X_{k+L}^R by X^R, having that $X^R \doteq \{X^{\neg I,U} \cup X^{I,U}\}$, with appropriate Jacobian partition $C^R \doteq \left[C^{\neg I,U}, C^{I,U}\right]$.

Next, noticing that $X_k = \{X_k^F \cup X^R\}$ we can partition the prior information matrix Λ_k respectively

$$\Lambda_k = \begin{bmatrix} \Lambda_k^F & \Lambda_k^{F/X^R} \\ (\Lambda_k^{F/X^R})^T & \Lambda_k^{X^R} \end{bmatrix}. \tag{C15}$$

Similarly, due to $X_{k+L}^R \doteq \{X^R \cup Y\}$ and $X_{k+L} = \{X_k^F \cup X^R \cup Y\} = \{X_k^F \cup X_{k+L}^R\}$ the posterior information matrix Λ_{k+L} can be respectively partitioned in next two forms:

$$\Lambda_{k+L} = \begin{bmatrix} \Lambda_{k+L}^F & \Lambda_{k+L}^{F/X^R} & \Lambda_{k+L}^{F/Y} \\ (\Lambda_{k+L}^{F/X^R})^T & \Lambda_{k+L}^{X^R} & \Lambda_{k+L}^{X^R/Y} \\ (\Lambda_{k+L}^{F/Y})^T & (\Lambda_{k+L}^{X^R/Y})^T & \Lambda_{k+L}^{Y} \end{bmatrix} = \begin{bmatrix} \Lambda_{k+L}^F & \Lambda_{k+L}^{F/X_{k+L}^R} \\ (\Lambda_{k+L}^{F/X_{k+L}^R})^T & \Lambda_{k+L}^{X_{k+L}^R} \end{bmatrix} \tag{C16}$$

with

$$\Lambda_{k+L}^{X_{k+L}^R} = \begin{bmatrix} \Lambda_{k+L}^{X^R} & \Lambda_{k+L}^{X^R/Y} \\ (\Lambda_{k+L}^{X^R/Y})^T & \Lambda_{k+L}^{Y} \end{bmatrix}. \tag{C17}$$

We can see from above partitions (C15)-(C17) that posterior information partition $\Lambda_{k+L}^{X_{k+L}^R}$ of X_{k+L}^R is simply the augmentation of prior information partition $\Lambda_k^{X^R}$ of X^R variables and can be calculated as:

$$\Lambda_{k+L}^{X_{k+L}^R} = \begin{bmatrix} \Lambda_k^{X^R} & 0 \\ 0 & 0 \end{bmatrix} + \begin{bmatrix} (C^R)^T \cdot C^R & (C^R)^T \cdot D \\ (D)^T \cdot C^R & (D)^T \cdot D \end{bmatrix} = \Lambda_k^{Aug,X^R} + A_R^T A_R \tag{C18}$$

where $A_R = [C^R, D]$, Λ_k^{Aug,X^R} can be constructed by first taking partition of prior information matrix Λ_k related to X^R, denoted by $\Lambda_k^{X^R}$, and augmenting it with n' zero rows and columns (see Figure 2c), where n' is just number of newly introduced variables.

The above equation has augmented determinant form as defined in Section III-A, and so the augmented determinant lemma can be applied also here. Using Eq. (10) we have:

$$\frac{\left|\Lambda_{k+L}^R\right|}{\left|\Lambda_k^{X^R}\right|} = \left|S\right| \cdot \left|D^T \cdot S^{-1} \cdot D\right| \tag{C19}$$

$$S = I_m + C^R \cdot (\Lambda_k^{X^R})^{-1} \cdot (C^R)^T \tag{C20}$$

Then by combining the Eq. (B10), Eq. (11) and the above equations, we can see that:

$$\frac{\left|\Sigma_{k+L}^F\right|}{\left|\Sigma_k^F\right|} = \frac{\left|\Lambda_{k+L}^R\right|}{\left|\Lambda_{k+L}\right|} \cdot \frac{\left|\Lambda_k\right|}{\left|\Lambda_k^{X^R}\right|} = \frac{\left|S\right| \cdot \left|D^T \cdot S^{-1} \cdot D\right|}{\left|P\right| \cdot \left|D^T \cdot P^{-1} \cdot D\right|} \tag{C21}$$

where P is defined in Eq. (13).

And apparently the IG of X_{k+L}^F can be calculated as:

$$J_{IG}^F(a) = \mathcal{H}(X_k^F) - \mathcal{H}(X_{k+L}^F) = \frac{1}{2}\ln\left|\Sigma_k^F\right| - \frac{1}{2}\ln\left|\Sigma_{k+L}^F\right| =$$
$$\frac{1}{2}(\ln\left|P\right| + \ln\left|D^T \cdot P^{-1} \cdot D\right| - \ln\left|S\right| - \ln\left|D^T \cdot S^{-1} \cdot D\right|), \tag{C22}$$

Next, S term can be further reduced. It is clear that $(\Lambda_k^{X^R})^{-1} = \Sigma_k^{X^R|F}$, or namely the prior conditional covariance matrix of X^R conditioned on X_k^F. Moreover, due to sparsity of C^R (its sub-block $C^{\neg I,U}$ contains only zeros) we will actually need only entries of matrix $\Sigma_k^{X^R|F}$ that belong to variables involved in new terms of Eq. (3) (see Figure 2c) and can conclude that:

$$S = I_m + C^R \cdot \Sigma_k^{X^R|F} \cdot (C^R)^T = I_m + C^{I,U} \cdot \Sigma_k^{X^{I,U}|F} \cdot (C^{I,U})^T \tag{C23}$$

■

Note that the variables inside information matrices do not have to be ordered in any particular way, and that the provided above proof is correct for any ordering whatsoever.

References

[1] D. Kopitkov and V. Indelman, "Computationally efficient belief space planning via augmented matrix determinant lemma and re-use of calculations," *IEEE Robotics and Automation Letters (RA-L)*, 2017, accepted.

[2] D. A. Harville, "Matrix algebra from a statistician's perspective," *Technometrics*, vol. 40, no. 2, pp. 164–164, 1998.

[3] D. Kopitkov and V. Indelman, "Computationally efficient decision making under uncertainty in high-dimensional state spaces," in *IEEE/RSJ Intl. Conf. on Intelligent Robots and Systems (IROS)*, October 2016.