UPPSALA UNIVERSITET

Preconditioning of Nonlinear Mixed Effect models for Stabilization of the Covariance Matrix Computation

Yasunori Aoki, Rikard Nordgren and **Andrew C. Hooker**

Department of Pharmaceutical Biosciences

Uppsala University

Sweden

UPPSALA UNIVERSITET

Motivating example

Estimation in NONMEM 7.3	OFV -2 ln (likelihood)	Estimated Parameter Value (V1)	Estimated SE (V1)
Linux Cluster Intel Xeon E5645 GCC 4.4.7	-2,346.78706	7.913632283	1.2618562800
MacBook Pro Intel Core i7 GCC 4.9.0	-2,346.78706	7.913632354	1.7525223004
Difference (%)	**0.000000001%**	**0.0000008939%**	**32.6%**

Siv Jönsson, Yi Fang Cheng, Charlotte Edenius, Kennedy R. Lees, Tomas Odergren, and Mats O. Karlsson. Population pharmacokinetic modelling and estimation of dosing strategy for NXY-059, a nitrone being developed for stroke. Clinical Pharmacokinetics, 44:863–878, 2005.

UPPSALA UNIVERSITET

Motivating example

Estimation in NONMEM 7.3	**OFV** -2 ln (likelihood)	**Estimated Parameter Value** (V1)	**Estimated SE** (V1)	**Estimated SE with preconditioning** (V1)
Linux Cluster Intel Xeon E5645 GCC 4.4.7	-2,346.78706	7.913632283	1.2618562800	1.38730205594
MacBook Pro Intel Core i7 GCC 4.9.0	-2,346.78706	7.913632354	1.7525223004	1.38466752752
Difference (%)	**0.000000001%**	**0.0000008939%**	**32.6%**	**0.190%**

Siv Jönsson, Yi Fang Cheng, Charlotte Edenius, Kennedy R. Lees, Tomas Odergren, and Mats O. Karlsson. Population pharmacokinetic modelling and estimation of dosing strategy for NXY-059, a nitrone being developed for stroke. Clinical Pharmacokinetics, 44:863–878, 2005.

UPPSALA UNIVERSITET

Why does this happen?

Sandwich estimator of the Covariance matrix of parameter estimates (M)

$$M = R^{-1} S R^{-1}$$

R: Hessian of the -2 ln likelihood

S: Sum of the cross products of the gradient vectors of the -2 ln individual-likelihood

Reminder: computation on a computer is not exact

– For example: machine epsilon gives an upper bound on the relative rounding error in floating point arithmetic (double precision ~ 10^{-15})

UPPSALA UNIVERSITET

Why does this happen?

Sandwich Estimator of the Covariance matrix

$$M = R^{-1} S R^{-1}$$

Computational error from rounding error $\lessapprox$ Condition number of R matrix

x

Small perturbation in S and R matrix

x

Condition number of R matrix

UPPSALA UNIVERSITET

Why does this happen?

Sandwich Estimator of the Covariance matrix

$$M = R^{-1} S R^{-1}$$

Computational error from rounding error $\lessapprox$ 4.623879×10^6 x Small perturbation in S and R matrix x 4.623879×10^6

UPPSALA UNIVERSITET

Why does this happen?

Sandwich Estimator of the Covariance matrix

$$M = R^{-1} S R^{-1}$$

Computational error from rounding error $\lessapprox$

4.623879×10^{6}

x

10^{-15}

x

4.623879×10^{6}

UPPSALA UNIVERSITET

Why does this happen?

Sandwich Estimator of the Covariance matrix

$$M = R^{-1} S R^{-1}$$

Computational error from rounding error	$\lessapprox$	10^{-3}
Computational error of SE from rounding error	$\lessapprox$	$\sqrt{10^{-3}} = 10^{-1.5}$

UPPSALA UNIVERSITET

How does preconditioning work?

Linearly re-parameterize the model by matrix P

$$\theta = P\hat{\theta}$$

to reduce the condition number of the $\hat{R}$ matrix

$$M = P\hat{R}^{-1}\hat{S}\hat{R}^{-1}P^{T} = P\hat{M}P^{T}$$

Computational error from rounding error	$\lessapprox$ Condition number of $\hat{R}$ matrix x Small perturbation in P, $\hat{S}$ and $\hat{R}$ matrix x Condition number of $\hat{R}$ matrix

UPPSALA UNIVERSITET

Choosing P

Assuming we are at the maximum likelihood with θ then we can obtain an eigen decomposition of R such that:

$$R = V \Lambda V^T$$

V : Normalized eigenvectors

Λ : Eigenvalues

then the $\hat{R}$ matrix condition number will be one if:

$$P = V \Lambda^{-1/2}$$

How does preconditioning work?

Linearly re-parameterize the model by matrix P

$$\theta = P\hat{\theta}$$

to reduce the condition number of the $\hat{R}$ matrix

$$M = P\hat{R}^{-1}\hat{S}\hat{R}^{-1}P^{T} = P\hat{M}P^{T}$$

Computational error from rounding error $\lessapprox$ 1.125299×10^{0}
x
Small perturbation in P, $\hat{S}$ and $\hat{R}$ matrix
x
1.125299×10^{0}

UPPSALA UNIVERSITET

Motivating example

Estimation in NONMEM 7.3	**OFV** -2 ln (likelihood)	**Estimated Parameter Value** (V1)	**Estimated SE** (V1)	**Estimated SE with preconditioning** (V1)
Linux Cluster Intel Xeon E5645 GCC 4.4.7	-2,346.78706	7.913632283	1.2618562800	1.38730205594
MacBook Pro Intel Core i7 GCC 4.9.0	-2,346.78706	7.913632354	1.7525223004	1.38466752752
Difference (%)	**0.000000001%**	**0.0000008939%**	**32.6%**	**0.190%**

Siv Jönsson, Yi Fang Cheng, Charlotte Edenius, Kennedy R. Lees, Tomas Odergren, and Mats O. Karlsson. Population pharmacokinetic modelling and estimation of dosing strategy for NXY-059, a nitrone being developed for stroke. Clinical Pharmacokinetics, 44:863–878, 2005.

UPPSALA UNIVERSITET

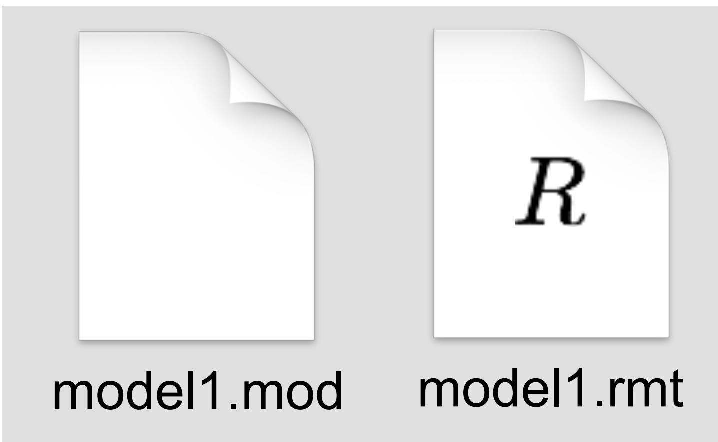
Preconditioning can be done using the **precond** command available in PsN 4.4

UPPSALA UNIVERSITET

precond model1.mod

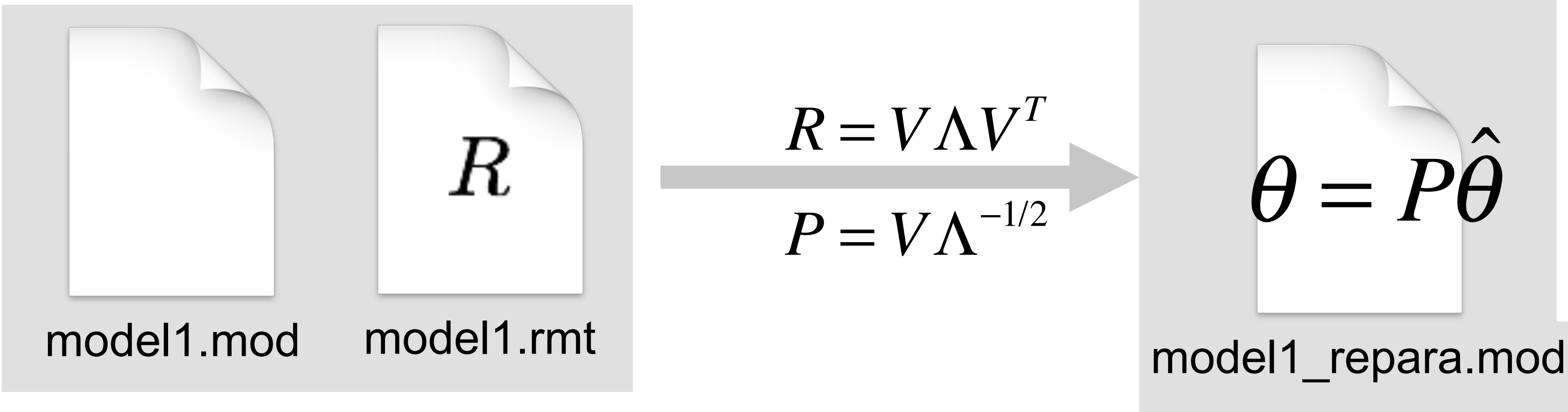
UPPSALA UNIVERSITET

precond model1.mod



UPPSALA
UNIVERSITET

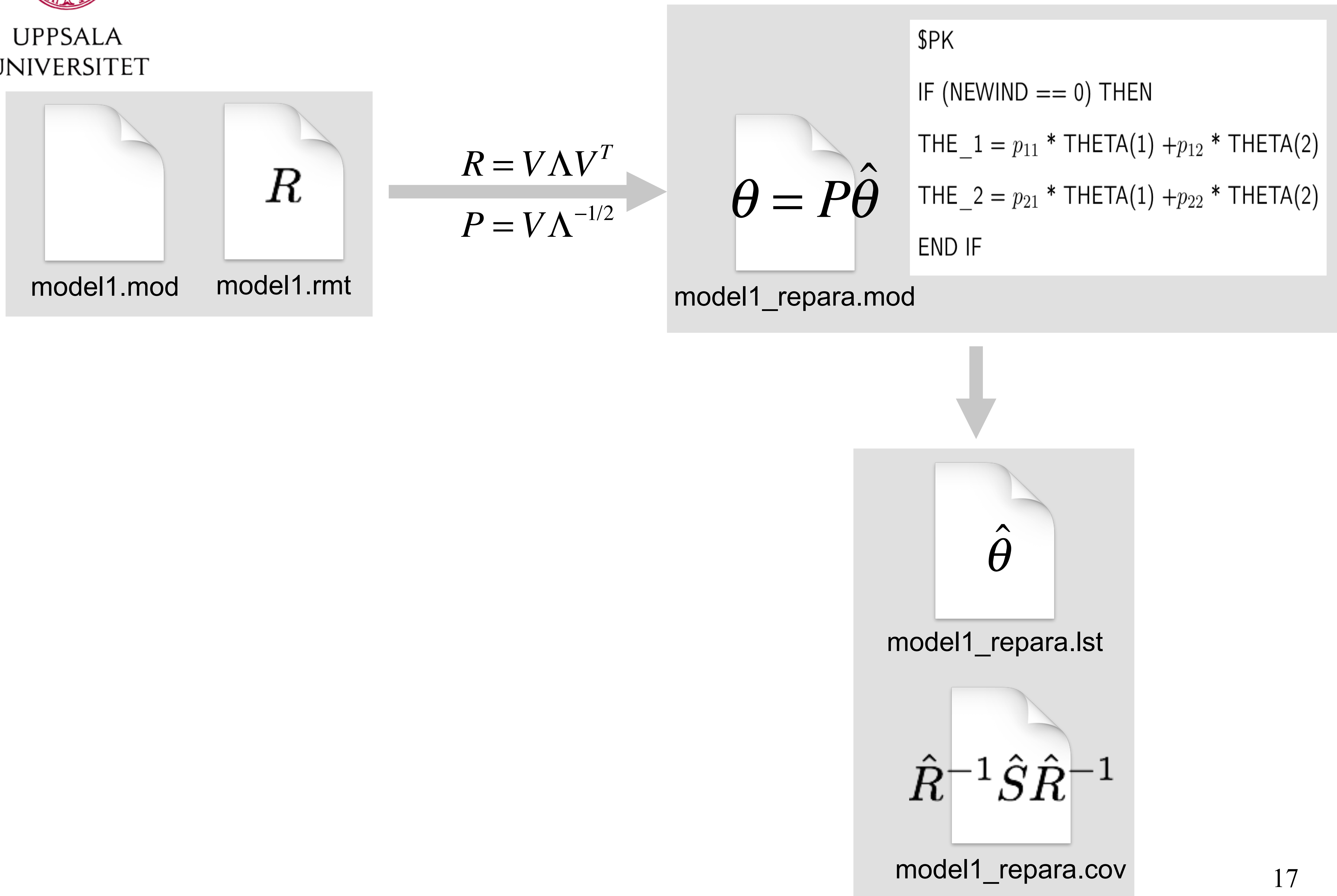
precond model1.mod




```
$PK
IF (NEWIND == 0) THEN
THE_1 = p11 * THETA(1) +p12 * THETA(2)
THE_2 = p21 * THETA(1) +p22 * THETA(2)
END IF
```

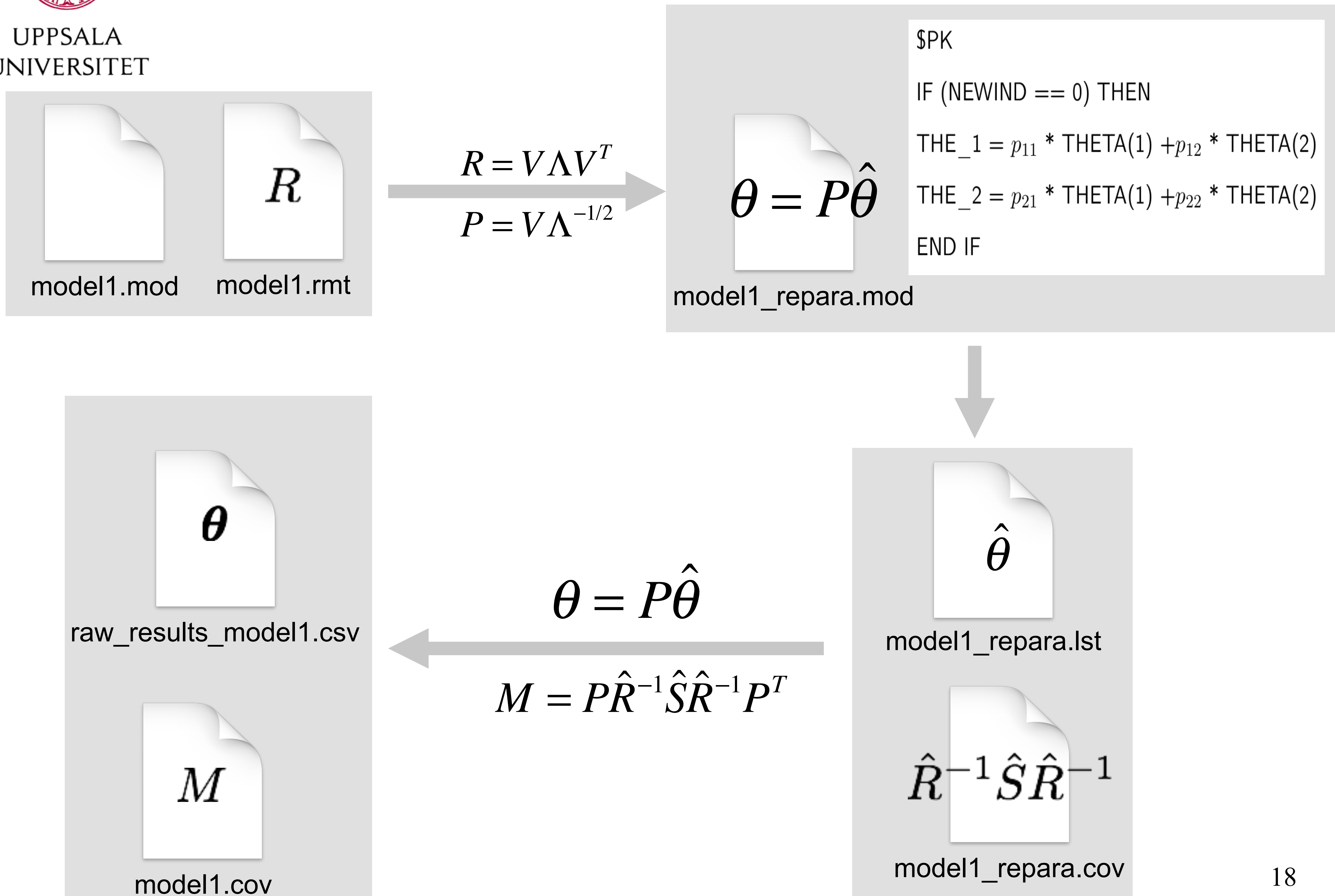
UPPSALA UNIVERSITET

precond model1.mod



precond model1.mod

UPPSALA UNIVERSITET



UPPSALA UNIVERSITET

Limitations with the ‘precond’ tool in PsN

- Neglects constraints on the parameter search space (i.e., boundaries of the fixed-effect parameters set in $THETA record).
  - Use “abs(THETA(X))” to account for non-negative parameter spaces.
- Only preconditions on the fixed effect portion of the model.
  - Re-parameterization of the model so that the standard deviation of random effects can be estimated with a fixed effect.
- Cannot precondition mixture models

UPPSALA UNIVERSITET

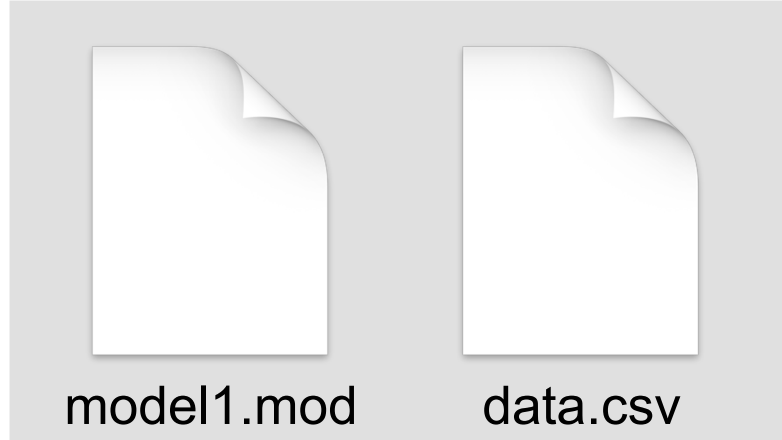
Preconditioning can

- Reduce computational environment dependencies
- Recover failed covariance matrix computations
- Aid in revealing model parameter non-identifiability

UPPSALA UNIVERSITET

Numerical Experiment 1

Reduce computational environment dependencies



Initial parameter estimates set to final parameter estimates from previous analysis

UPPSALA
UNIVERSITET

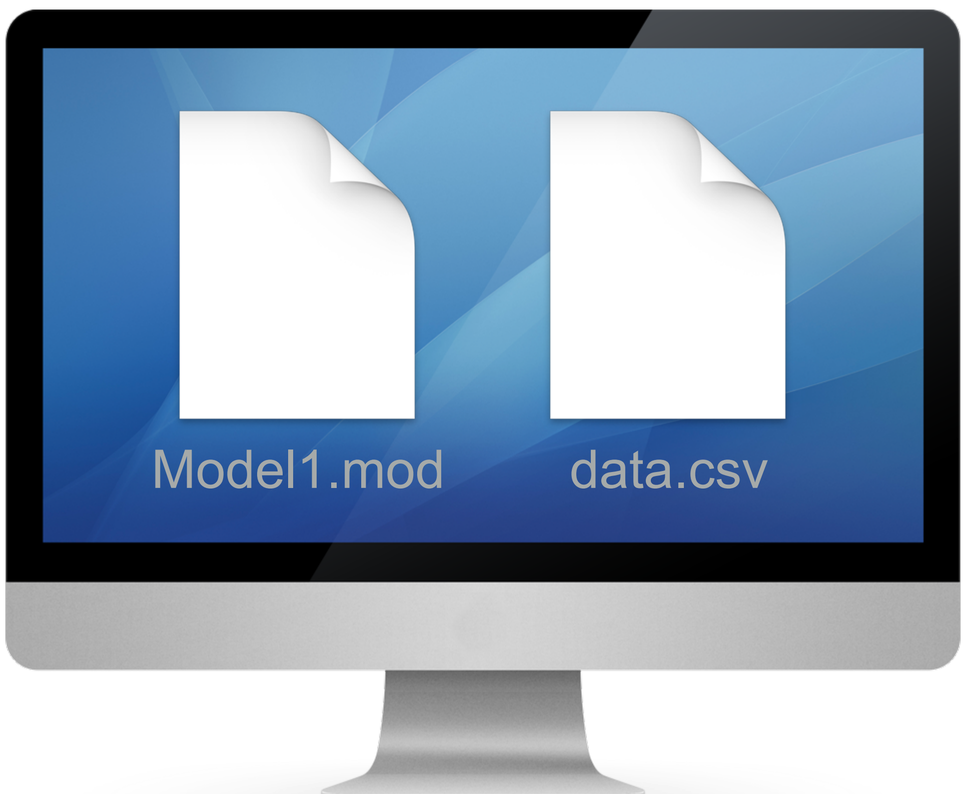
model1.mod data.csv

Initial parameter estimates set to final parameter estimates from previous analysis

Linux Cluster

MacBook Pro

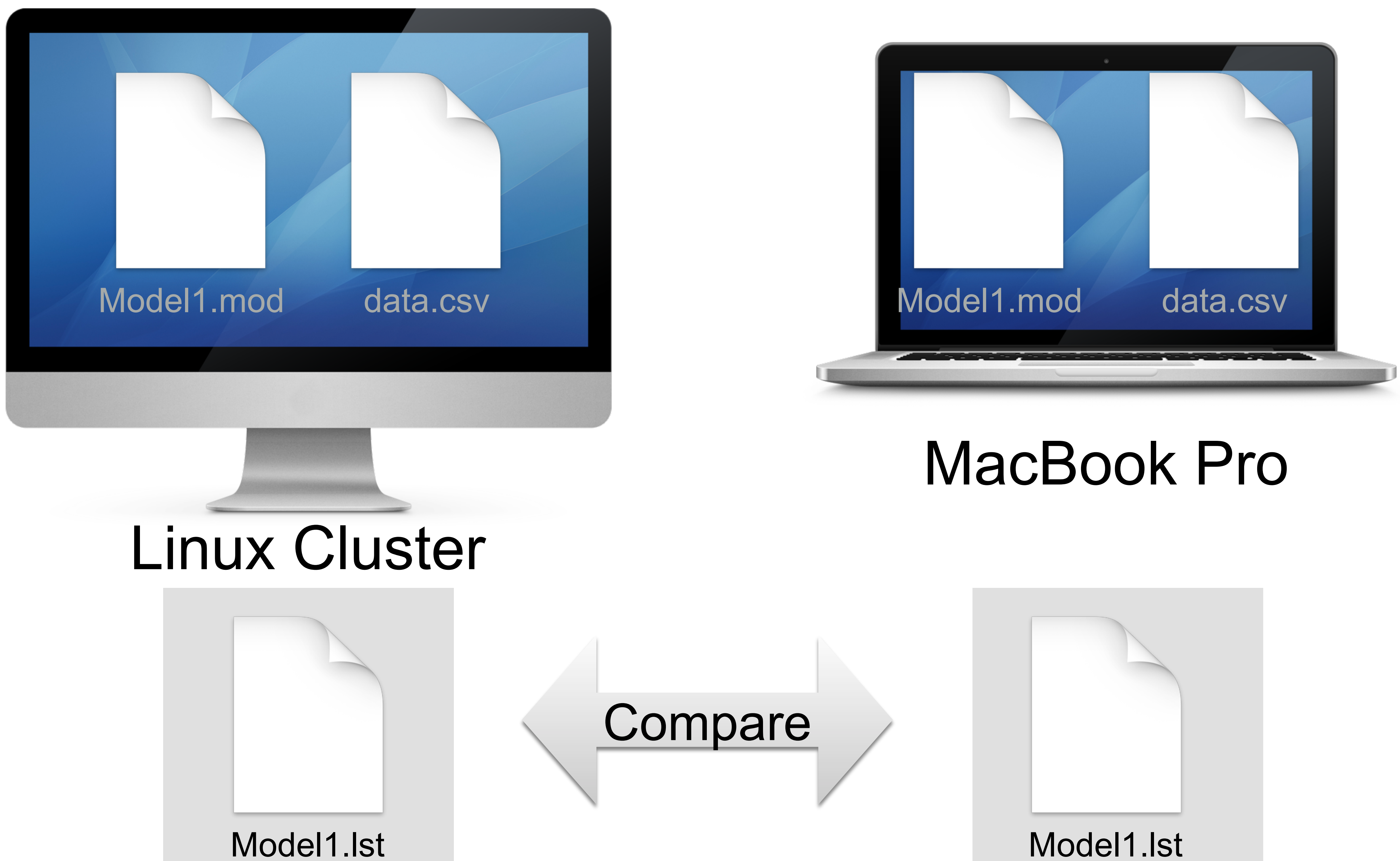
UPPSALA
UNIVERSITET



Linux Cluster

MacBook Pro

UPPSALA
UNIVERSITET



UPPSALA UNIVERSITET

Original model results

	R Condition Number	Difference in OFV	Ave. Difference of SE (%)
Model1 [1]	4.623879×10^{6}	0.000000028020	21.2%
Model2 [2]	3.674548×10^{11}	0.000000203632	5.47%
Model3 [3]	4.795944×10^{6}	0.000000858280	11.2%

[1] Jönsson, et al., Clinical Pharmacokinetics, 44:863–878, 2005.
[2] Bergmann, et al., *The pharmacogenomics journal*, 11:113–120, 2011.
[3] Wählby, et al., British Journal of Clinical Pharmacology, 58:367–377, 2004.

UPPSALA UNIVERSITET

Preconditioned model results

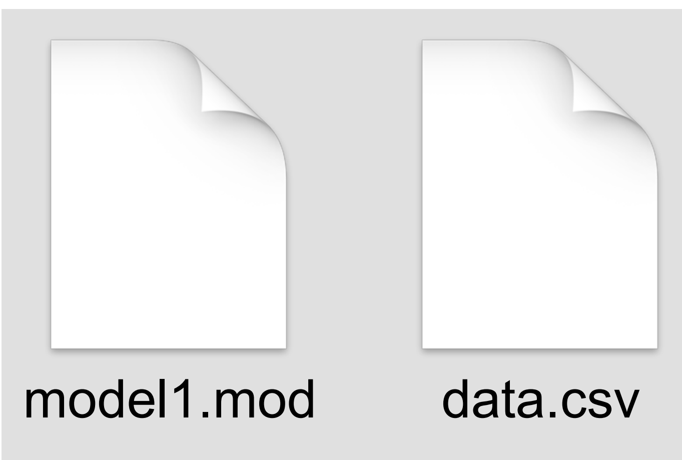
	R Condition Number	Difference in OFV	Ave. Difference of SE (%)
Model1 [1]	4.623879×10^{6}	0.000000028020	21.2%
preconditioned model1	**1.125299×10^{0}**	**0.000000140300**	**0.101%**
Model2 [2]	3.674548×10^{11}	0.000000203632	5.47%
preconditioned model2	**2.630800×10^{2}**	**0.000000071808**	**0.504%**
Model3 [3]	4.795944×10^{6}	0.000000858280	11.2%
preconditioned model3	**1.533025×10^{0}**	**0.000000640840**	**0.297%**

[1] Jönsson, et al., Clinical Pharmacokinetics, 44:863–878, 2005.
[2] Bergmann, et al., *The pharmacogenomics journal*, 11:113–120, 2011.
[3] Wählby, et al., British Journal of Clinical Pharmacology, 58:367–377, 2004.

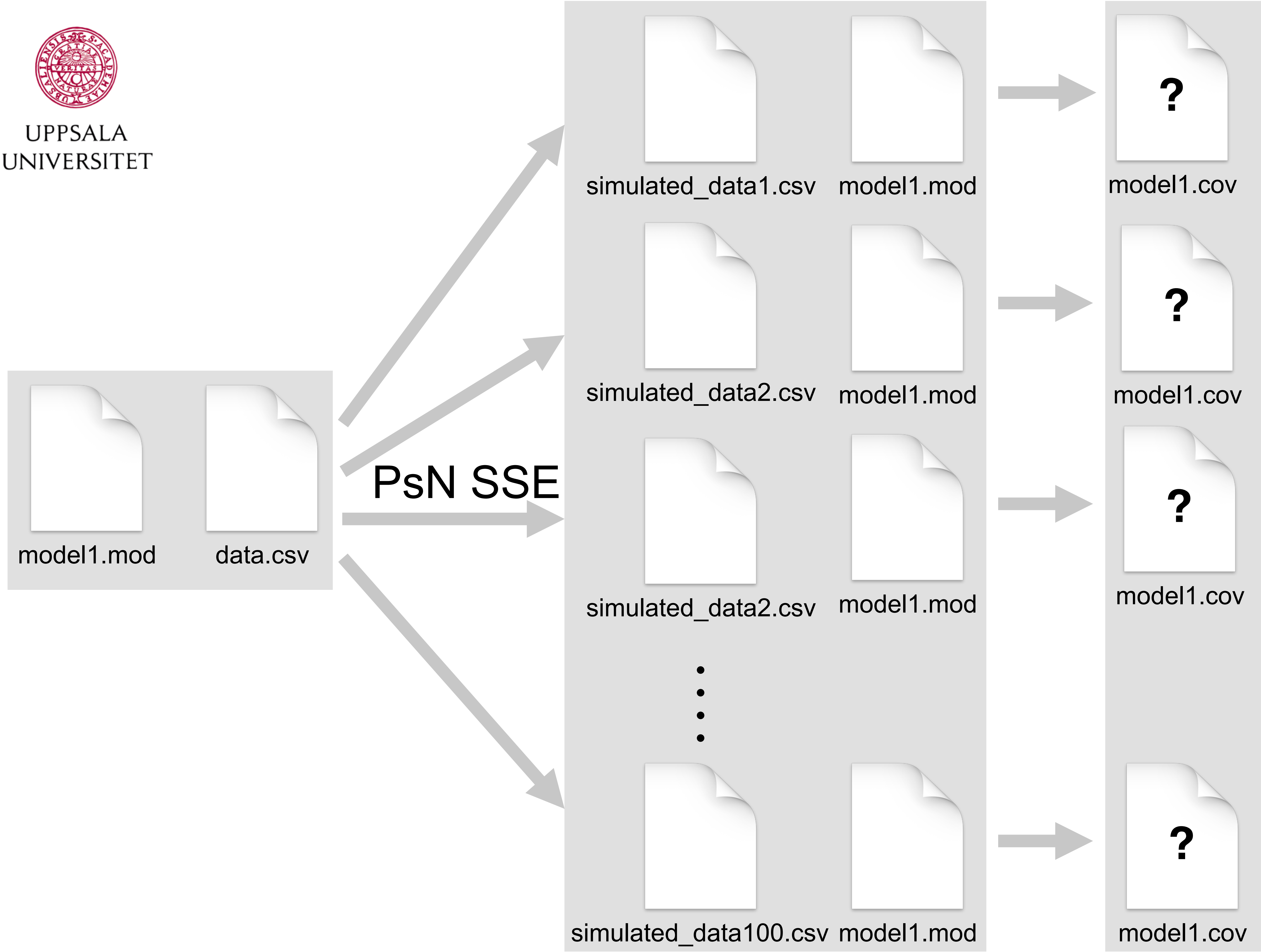
UPPSALA
UNIVERSITET

Numerical Experiment 2

Recover failed variance-covariance matrix computations


model1.mod
data.csv

UPPSALA UNIVERSITET



UPPSALA
UNIVERSITET

Results

	Original Model
Model 1	100/100
Model 2	68/100
Model 3	62/100

UPPSALA UNIVERSITET

Results

	Original Model	Preconditioned Model
Model 1	100/100	100/100
Model 2	68/100	100/100
Model 3	62/100	98/100* **

*Estimation of the original model failed
**S-matrix not obtained for original or preconditioned model

UPPSALA
UNIVERSITET

Numerical Experiment 3

Aid in revealing model parameter non-identifiability

UPPSALA UNIVERSITET

Original model results [4]

	OFV	CL	V1	Q2	V2	Q3	V3	RUV
Original model	186.145	0.36	3.44	1.94	0.69	1.22	3.25	0.045

No covariance step:

"R MATRIX ALGORITHMICALLY SINGULAR AND ALGORITHMICALLY NON-POSITIVE-SEMIDEFINITE”

[4] Alskär et al., J. PKPD, 39:273-282, 2012

UPPSALA UNIVERSITET

Parameters obtained using preconditioning

	OFV	CL	V1	Q2	V2	Q3	V3	RUV
Original model	186.145	0.36	3.44	1.94	0.69	1.22	3.25	0.045
Preconditioned model	186.145	19.36	187.32	105.53	37.34	66.70	177.10	0.045

- Same OFV, very different parameter values
  - Used the parameters obtained with preconditioning as the initial estimate of the original model and obtained the same OFV.
- The smallest eigenvalue of the preconditioned R-matrix is -5.95×10^{-10} indicating R-matrix to be singular.
- Hence this model is highly likely to be unidentifiable.

UPPSALA UNIVERSITET

SSE Study

- In 27/100 cases variance-covariance matrix was obtainable for the original model and RSE of all parameters were less than 50%.
- Typical RSE with preconditioning were orders of magnitude larger:

	CL	V1	Q2	V2	Q3	V3	PropErr.
Original RSE	7.68%	6.49%	21.87%	23.73%	31.99%	25.93%	5.78%
RSE with Preconditioning	4024%	4027%	4124%	4089%	3985%	4069%	5.93%

UPPSALA UNIVERSITET

Preconditioning can

- Reduce computational environment dependencies
- Recover failed covariance matrix computations
- Aid in revealing model parameter non-identifiability

- **precond** available in PsN 4.4 (psn.sf.net)

- Computational instability can also influence the parameter estimates and an investigation of this correlation using the preconditioning method is presented as a poster:
  - Bjugård Nyberg *et al.* Influence of Covariance Step Success on Final Parameter Estimates. PAGE 24 (2015) Abstr 3601 [www.page-meeting.org/?abstract=3601].

Acknowledgement: This work was supported by the DDMoRe (www.ddmore.eu) project